

Alkalmazott matematikai lapok

1983/1-2

A MAGYAR TUDOMÁNYOS AKADEμία
MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

9

KÖTET

AKADÉMIAI KIADÓ, BUDAPEST

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI
TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

FŐSZERKESZTŐ

PRÉKOPA ANDRÁS

FŐSZERKESZTŐ-HELYETTES

ARATÓ MÁTYÁS

A SZERKESZTŐ BIZOTTSÁG TAGJAI

BENCZUR ANDRÁS, CSISZÁR IMRE, FARKAS MIKLÓS, GYIRES BÉLA,
HATVANI LÁSZLÓ, HEPPES ALADÁR, KÁTAI IMRE, KIS OTTÓ,
SARKADI KÁROLY, TANDORI KÁROLY, VARGA LÁSZLÓ,
SZÁNTAI TAMÁS (TECHNIKAI SZERKESZTŐ)

MUNKATÁRSAK

BAJCSAY PÁL, BALLA KATALIN, BÉKÉSSY ANDRÁS, CSÁKI PÉTER,
CSIRIK JÁNOS, DEMETROVICS JÁNOS, DÉNES JÓZSEF, DÖMÖLKI BALINT,
ELBERT ÁRPÁD, FORGÓ FERENC, GÉCSEG FERENC, GERGELY JÓZSEF,
GESZTELYI ERNŐ, GYÖRFFY LÁSZLÓ, KLAFSZKY EMIL, KÖSA ANDRÁS,
KOVÁCS LÁSZLÓ BÉLA, LÁSZLÓ ZOLTÁN, MIKOLÁS MIKLÓS,
MOGYORÓDI JÓZSEF, NÉMETH GÉZA, NEMETZ TIBOR, RÉVÉSZ PÁL,
RÓZSA PÁL, STAHL JÁNOS, SZÉP JENŐ, TANKÓ JÓZSEF, TOMKÓ JÓZSEF,
TÓKE PÁL, TUSNÁDY GÁBOR, VINCZE ENDRE

IX. kötet 1—2. szám

Szerkesztőség: 1502 Budapest XI., Kende u. 13—17.

Kiadóhivatal: 1055 Budapest V., Alkotmány u. 21.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztő bizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Prékopa András, főszerkesztő
1502 Budapest, Kende u. 13—17.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 100 forint. Belföldi megrendelések az Akadémiai Kiadó, 1055 Budapest V., Alkotmány u. 21. címen (pénzforgalmi jelzőszám 215—11 488), külföldi megrendelések a Kultúra Külkereskedelmi Vállalat, H-1389 Budapest, Pf. 149. címen (pénzforgalmi jelzőszám 218—10 990) lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungaricae,
2. Acta Physica Hungaricae,
3. Studia Scientiarum Mathematicarum Hungarica.

MATEMATIKA ÉS OBJEKTÍV VALÓSÁG

FARKAS MIKLÓS

Budapest

1. Bevezetés

A matematikai fogalmak viszonya az objektív valósághoz foglalkoztatta már az ókori filozófusokat is. A 19. századot megelőző korokban azonban ez a kérdés meglehetősen ritkán merült fel és csupán mellékes része, vetülete volt annak az általános filozófiai vitának és küzdelemnek, amely az anyagi világ, illetve az emberi tudat viszonya, elsődlegessége körül folyt materialista és idealista filozófusok között. Ennek okát abban látom, hogy egészen a 19. századig a matematika alig nevezhető önálló tudománynak. A matematika a földmérés, a csillagászat, a mechanika, a navigáció és a fizika segédeszközeként, ezekkel a tudományokkal szorosan összefonódva fejlődött ki. Általános, jelentős matematikai tételek és fogalmak a szóban forgó korszakban gyakran nem ma megszokott absztrakt formájukban jelentek meg, hanem egészen konkrét, a mindennapi életből vett jelenségekhez kapcsolt alakban.

Ennek a korszaknak nem igen van olyan jelentős tudósa, aki kizárólag matematikával foglalkozott volna. E korszak természettudósai személyükben is egyesítették a matematikát a mechanikával, a csillagászattal, a fizikával, sőt a fiziológiával és a művészetekkel is. A tudomány- és művészettörténet nagy alakjait általában aszerint soroljuk be az egyik vagy másik tudománykategóriába, hogy milyen téren alkotnak többet, maradandóbbat. Ezek a besorolások azonban óhatatlanul erőltetettek és szubjektívek. Megtehetjük például azt, hogy FERMAT-t matematikusnak, EULERT és a BERNOULLI család legtöbb ismert tagját matematikusnak, NEWTONT elsősorban mechanikusnak tekintsük. Pedig, ha FERMATTól semmi nem maradt volna ránk, csupán az optika *Fermat-elve*, EULERTól és a BERNOULLIaktól csupán a hidrodinamika alapegyenletei és a NEWTONTól csak a differenciálszámítás, nevük akkor is a tudománytörténet nagyjai között sorakozna.

Körülbelül a 19. század elejét tekinthetjük annak a kornak, amikor a műszaki haladás és a tudományok fejlődése következtében az egyes tudományok élesen elválnak egymástól. A korszakváltás, a matematika élesen elhatárolt, külön tudománnyá válása e tudomány addig nem tapasztalt mértékű fejlődését és kiterjedéyesedését követte. Olyan nevek jellemzik e korszakot, mint a nagy franciáké (LAPLACE, CAUCHY, FOURIER stb.), mint GAUSSÉ, illetve BOLYAI JÁNOSÉ. Ez utóbbi nevet nem csupán a nemzeti büszkeség íratja ide, hanem az, hogy éppen BOLYAIval indult útjára a modern matematikában az axiomatikus módszer. Márpedig az axiomatikus módszer térhódítása és a matematikai szigorúság minőségileg magasabb szintre emelése alapvető szerepet játszott a matematika leválásában a többi természettudományról.

A matematika önálló, külön tudománnyá válása tovább gyorsította fejlődését, elvezetett e tudomány sajátos nyelvezetének, módszereinek kikristályosodásához. Megjelentek a csak matematikával foglalkozó tudományos kutatók, és megszületett a *tiszta matematika*. Ennek a fejlődésnek következtében egyre nehezebbé vált a válaszadás arra a kérdésre, hogy mi is a matematika tárgya, és mi a matematikai fogalmak, eredmények viszonya az objektív valósághoz. Az elmúlt évtizedekben hazánkban is többen foglalkoztak ezzel a kérdéssel és a legkülönbözőbb oldalokról közelítették meg a választ. A témával foglalkozó, tudományos folyóiratokban megjelent dolgozatok egy része kétségbe vonja azt, hogy a matematika tudományának tárgya az objektív valóság, kategorikusan tagadja azt, hogy a matematika bármilyen mértékben és értelemben is természettudománynak lenne tekinthető, tagadja azt, hogy az alkalmazott matematika matematika és végső fokon misztifikálja ezt a tudományt.

Más szerzők kísérletet tesznek arra, hogy a matematikát részben a művészetek közé sorolják be, figyelmen kívül hagyva a tudományos, illetve a művészi visszatükrözés lényeges különbségeit.

2. A matematikai fogalmak viszonya az objektív valósághoz

Minden tudomány absztrakt fogalmakkal dolgozik. Az emberi tudat képtelen a valóságot teljes összetettségében, egyszerre megragadni és visszatükrözni. Kénytelen elhanyagolni lényegtelen vonásokat, eltekinteni a mellékes jelenségektől és kiemelni a lényegyet, illetve a vizsgálat szempontjából lényeges tulajdonságokat. Éppen ez a képesség, ti. az elvonatkoztatás a jelenlevő, de lényegtelen jegyeiktől, kísérő jelenségektől és a lényeges vonások megtalálása, az emberi gondolkodás egyik legerősebb fegyvere. Elefánt nincs, csak elefántok vannak, hiszen minden elefánt különbözik az összes többi elefánttól genotípusában, és ha mégis léteznének egypetéjű elefántikrek, fenotípusuk akkor is megkülönböztethető lenne. Tiszta oxigén gáz nem létezik, mert mindig jelen van bizonyos egyéb molekulák által okozott szennyezettség. A Föld nem kering a Nap körül, hanem mindkét égitest közös súlypontjuk körül kering és még egyéb bonyolult mozgásokat is végez. Ennek ellenére a zoológus releváns, minden elefántra érvényes dogákat képes mondani erről az állatfajról, és minden gyerek felismeri az elefántot az állatkertben; a vegyész tudja, hogy mi történik, ha a hidrogént tiszta oxigén jelenlétében meggyújtja, és a valóságban, valóságos, szennyezett oxigéngáz jelenlétében lényegében az történik, amit a vegyész várt; ha időszámításunkat és mondjuk az állócsillagok megfigyelését arra a tudományos megállapításra alapozzuk, hogy a Föld 365 és néhány tized nap alatt kerüli meg a Napot, értékes, pontos eredményeket kapunk.

Az előbb azt írtam, hogy elefánt nincs; a valóságban azonban objektíve léteznek azok a lényeges tulajdonságok, amelyek alapján eldönthetjük egy élőlényről, hogy elefánt-e vagy sem. Ezek a tulajdonságok a biológia tudományától, az emberi tudattól függetlenül léteznek olyannyira, hogy az az egészséges elefántbika, amelyik még sohasem látott biológust, sem próbálkozik egy zebra kancával utódot nemzeni. Az az absztrakció, hogy elefánt, tehát bizonyos a valóságban objektíve létező tulajdonságok összességének, az „elefántnak lenni” összes sine qua non-jának tükröképe az emberi tudatban.

Minden tudomány alkalmaz több lépcsős absztrakciókat. Ilyen lépcsők például a biológiában: az elefánt, az emlős, a gerinces, az állat, az élőlény; a fizikában első

lépcsőként a hővezetés, a folyadékáramlás, illetve az impulzusáram, második lépcsőként pedig a transport-folyamat fogalma, amely esetében közömbös, hogy hő, tömeg, impulzus, vagy valami egyéb áramlik-e.

A matematika minden más tudománynál erőteljesebben absztrahál és minden más tudománynál több lépcsős absztrakciókkal dolgozik. A matematikai fogalmak azonban éppúgy a valóság objektív tulajdonságainak, viszonyainak tükröképei a tudatban, mint a többi tudomány fogalmai.

A természetes számok az emberiség legősibb matematikai absztrakciói. A hatos szám kifejezi mindazon véges halmazok közös tulajdonságát, amelyeknek hat elemük van. A hatos szám objektíve nem létezik az anyagi valóságban, ez a fogalom az emberi tudat terméke, ha úgy tetszik, pusztá forma (ugyanúgy, mint az elefá-fogalma). A hat elemű halmazoknak az a tulajdonsága, hogy pontosan hat elemük van, azonban objektív, valóságos tulajdonság, amely megkülönbözteti őket például az öt- vagy hételemű halmazoktól. Amikor a matematikus a természetes számokról megállapít valamilyen törvényszerűséget, akkor egyben az összes véges sok elemű halmaz objektív tulajdonságáról, a tudatunktól függetlenül létező valóságról mond valamit. Az az állítás például, hogy a hatos szám páros, kifejezi minden egyes konkrét, valóságos hatelemű halmaznak azt a tulajdonságát, hogy két egyenlő részre osztható.

Hasonló módon a geometriák axiómái és primitív fogalmai eredetileg a valóságos, fizikai térben fennálló, vagy lehetségesnek gondolt viszonyokból absztrahálódtak. A háromdimenziós euklideszi téergeometriának az az axiómája, amely szerint három nem egy egyenesen levő ponton át pontosan egy sík fektethető, kifejezi azt a valóságosan tapasztalt jelenséget is, hogy háromlábú asztal sohasem billeg, négylábú viszont igen gyakran.

Az a megállapítás, hogy a 2π hosszúságú intervallumon négyzetesen integrálható függvények Hilbert-teret alkotnak és e térnek a $\cos kx$, $\sin kx$ függvényrendszer bázisát képezi, egyben kifejezi a valóságos, kifeszített hegedűhúrnak azt a tulajdonságát, hogy bármilyen reálisan megvalósítható állapota előállítható az előbbi függvények által leírt állapotok superpozíciójaként, vagyis, hogy elegendő e húr „harmonikus rezgéseivel” foglalkoznunk.

A hasonló példákat vég nélkül lehetne sorolni úgy, hogy általában egyetlen matematikai megállapítás számos egymástól különböző valóságos jelenségről mondjon valami fontosat. Az erős absztrakciónak a matematikában éppen az az egyik értelme, hogy ezáltal válik egy-egy matematikai elmélet, illetve modell szerteágazó, széles területeken alkalmazhatóvá.

Mindebből pedig következik az, hogy miben különbözik a matematika a saktól, a bridzstől, a kép- vagy keresztrejtvényfejtéstől, általában a szellemi játékoktól. A szellemi játékokban az emberi goldokodás által létrehozott szabályok és konvenciók keretei között összefüggéseket keresünk a „pusztá formák” között; a helyes stratégia alkalmazása, az összefüggések megtalálása a játék megnyerésére, illetve a rejtvény megoldására vezet. A játék fontos dolog az ember, a homo ludens életében, azonban az optimális lépés megtalálása a sakkjátszmában, a helyes húzás felismerése a bridzsbén, a rejtvény megoldása semmivel sem visz közelebb az objektív valóság megismeréséhez, nem is ez a célja.

A matematikai szabályok, axiómarendszerek, elméletek ugyanúgy az emberi goldokodás alkotásai, mint a játékszabályok. *A matematikában azonban az elméletek a meghatározott, absztrakt szemszögből nézett objektív valóság modelljei, tükröképei az emberi tudatban. Ezért minden helytálló matematikai eredmény, felismerés*

nem csupán az absztrakt fogalmak, gondolkodási formák, szimbólumok közötti összefüggés megállapítása, hanem egyben mindazon valóságos jelenség valamilyen valóságos tulajdonságának megállapítása is, amely a szóban forgó fogalmakkal, szimbólumokkal modellezhető.

A tudománytörténet számos olyan esetet ismer, amikor a tudomány belső fejlődéstörvényei parancsának engedelmeskedve, az általánosítás lehetőségét kihasználva matematikusok absztrakt elméleteket hoztak létre, és jóval később derült ki, hogy ezek az anyagi világ természettudományos leírásának alkalmas matematikai modelljei.

A kúpszeletek elméletét időszámításunk előtt a harmadik században alkották meg. Az akkor ismert eredmények összefoglalása APOLLONIOSZ könyveiben található, amelyek az i. e. harmadik század végén — második század elején íródtak. Az első igazán fontos alkalmazások KEPLER törvényei a 17. században.

B. RIEMANN 1854-es magántanári előadásában GAUSS felületelméletét általánosítva alkotta meg a *Riemann-geometriát*. Záró megjegyzéseiben mentegetődzik a haszontalan téma miatt. Több mint ötven évvel később ismerte fel EINSTEIN, hogy éppen a RIEMANN által konstruált geometria alkalmas az általános relativitáselmélet tér-idejének modelljéül.

A csoportelmélet a múlt század első felében, az algebrai egyenletek megoldhatóságának vizsgálata során fejlődött ki és vált önálló matematikai diszciplínává. Majdnem egy évszázaddal később, a kvantummechanika megszületése után derült ki, hogy a csoportelmélet nélkülözhetetlen eszköz a mikrofizikai jelenségek leírásához.

A topológia és különösen az algebrai topológia a modern matematika egyik legabsztraktabb fejezete, amely legalábbis az elmúlt 100 évben szinte kizárólag a tudomány belső fejlődéstörvényeit követve fejlődött. Ez az absztrakt tudományterület szülte meg a *René Thom-féle elemi katasztrófaelméletet* a hetvenes évek elejére. Az utóbbi elmélet a potenciális dinamikai rendszerekkel modellezhető és elég kevés külső, vagy kontrollparaméterrel rendelkező jelenségek, folyamatok ugrásszerű változásainak direkt matematikai modellje.

Látjuk tehát, hogy az időnként hangoztatott állítás, mely szerint a matematika csupán „puszta formákkal” foglalkozik és eredményei semmit sem mondanak a valóságos dolgokról, felületes megállapítás. Ilyen alapon azt is mondhatnánk, hogy pl. a biológia is „puszta formák” között állapít meg összefüggéseket, amikor mondjuk, az emlősökről állít valamit, hiszen az „emlős” fogalom absztrakció. Nyilvánvaló azonban, hogy egy az emlősökre vonatkozó megállapítás mond valamit minden egyes a valóságban létező emlősről is.

A matematikai absztrakciók kapcsolatát a valósággal természetesen nehezebb felismerni, mint a biológiai, kémiai, fizikai, közgazdasági stb. absztrakciókét. Ez utóbbiak ugyanis mind a valóság egy jól körülhatárolt szeletére vonatkoznak, míg az előbbiek minden olyan valóságos jelenségre, amelynek bizonyos specifikus tulajdonsága a szóban forgó matematikai absztrakcióval modellezhető. *Azonban abból, hogy ez a kapcsolat rejtettebb és bonyolultabb, nem következik az, hogy nem is létezik.*

A több lépcsős matematikai absztrakciók és a valóság viszonyának megértéséhez foglalkoznunk kell egy keveset magának az objektív valóságnak fogalmával is. Az emberi tudat különböző módokon tükrözi a valóságot és a gondolkodáson alapuló emberi cselekvés visszahat a világra. Az objektív valóság és annak tudati tükörképe lényegesen különböző kategóriák, azonban a kapcsolat közöttük dialektikus. Ennek a dialektikus kapcsolatnak csak egyetlen vonásával foglalkozom itt, éspedig

a következővel: különböző nézőpontokból és vonatkozásaiban vizsgálva egy tudati tükörképet *relatívvá válhat az, hogy a szóban forgó dolgot objektív realitásnak, vagy az emberi tudat által létrehozott tükörképnek kell-e tekintenünk.*

LEV TOLSZTOJ hatalmas történelmi regénye, a „Háború és béke” utolérhetetlen módon ötvözi az általános érvényű, minden kor minden emberéhez szóló mondani-valót a korszak orosz viszonyainak hűséges ábrázolásával. Ez a könyv tehát minden egyéb nagyszerű vonása mellett, amelyek a regényirodalom egyik legnagyobb, vagy talán a legnagyobb alkotásává teszik, a korabeli orosz társadalom és ember magas színvonalú, művészi visszatükrözése. Minden történésznek, aki a napóleoni háborúk korszakát vizsgálja, ismernie kell ezt a regényt, de figyelembe kell vennie azt is, hogy ez tudati tükörkép, egy LEV TOLSZTOJ nevű, a 19. század második felében tevékenykedő író alkotása, és az ábrázolt korszakot a szerző gondolkodásmódjában, személyes tapasztalataiban, érzelmeiben és filozófiájában megfürdetve tükrözi. Ugyanakkor egy irodalomtörténész, vagy filológus a létező regényt objektív valóságnak tekintheti. Az irodalomtörténész, vagy filológus is tisztában van természetesen azzal, hogy a regény azt tartalmazza, ahogyan TOLSZTOJ tudata a 19. század második felében visszatükrözte a 19. század elejének orosz valóságát. Ha azonban őt első-sorban TOLSZTOJ élete, gondolkodása, művészi módszere, nyelvezete, mondatszerkesztése stb. érdekli, akkor a regényt, mint a tudatunktól függetlenül létező, objektív valóság részét tanulmányozza.

Ezzel a példával azt a véleményemet kívántam alátámasztani és megvilágítani, amely szerint *a valóság művészi vagy tudományos visszatükrözése, ha már létrejött és rögzítődött* (művészeti vagy tudományos alkotás formájában), *bizonyos összefüggésekben és vonatkozásokban maga is az objektív valóság részévé válik.*

Ez a megállapítás hozzásegít ahhoz, hogy jobban megértsük a több lépcsős matematikai absztrakciók létrejöttének dinamikáját és az absztrakt matematikai fogalmak áttételes viszonyát az objektív valósághoz.

A matematikus kutató munkája közben az általa tanulmányozott elmélet axiómáit, fogalmait, már feltárt összefüggéseit, tételeit objektíve létezőknek tekinti. A matematikai objektumok egyrészt a külső valóság bizonyos vonásainak tudományos tükörképei, másrészt létrejöttük után, az őket tanulmányozó számára objektíve adottak, az emberi tudat által alkotott objektív realitások.

Bizonyos értelemben ez a helyzet a már megalkotott elméletek még meg nem alkotott tételeivel is. Ha ugyanis valaki összeállítja az euklideszi geometria axiómáit, akkor az axiómák által generált elméletben objektíve ott rejtőzik PITHAGORASZ tétele is, akár felfedezte ezt már PITHAGORASZ (vagy valaki ő előtte), akár nem. Ha a valós számtestet definiáló axiómarendszer adott, akkor a valós számhalmazt önmagába leképező függvények elméletében objektíve jelen van az a tulajdonság, amelyet BOLZANO tételének nevezünk, ti. az, hogy intervallumon folytonos függvény két értéke között minden értéket felvesz. *A tételek ott rejtőzködnek az elméletek mélyén és várják, hogy felfedezzék őket.* Ugyanúgy, ahogyan Amerika ott rejtőzött az Atlanti óceánon túl és várta, hogy KOLUMBUSZ felfedezze, amajd AMERIGO VESPUCCI új kontinens voltát megállapítsa.

A több lépcsős matematikai absztrakció kissé leegyszerűsített sémája a következő. A környező világ bizonyos jelenségeinek tudományos vizsgálata eredményeképpen létrejön egy *fizikai, biológiai, közgazdasági* stb. megállapítás, tétel, elmélet, közös néven *modell*. Ilyen például a heliocentrikus naprendszer, az energiamegmáradás törvénye, a természetes kiválasztódás elve, az értéktöbblet-elmélet, az a megállapí-

tás, hogy „az emberi társadalom története osztályharcok története” stb. E modelleket objektíve létezőknek tekintjük és így tanulmányozzuk. Bizonyos elvonatkoztatás után a modellek lényeges vonatkozásait egzakt formában tükröző *matematikai modelleket* alkotunk. E matematikai modelleket mint objektív realitásokat tanulmányozzuk, kiemeljük ilyen vagy olyan szempontból lényeges tulajdonságait, illetve több különböző modell közös tulajdonságait és ezekből az elvonatkoztatás lépcsőjén feljebb lépve újabb, absztraktabb matematikai modellt hozunk létre. És így tovább.

A valóságos testek tulajdonságait, a minket körülvevő „fizikai teret” vizsgálva, ennek első matematikai modelljeként létrejött az euklideszi tér fogalma. Az euklideszi geometriát tanulmányozva kiderült, hogy ha a metrikus tulajdonságoktól eltekintünk, de a párhuzamosság tulajdonságát és a párhuzamos szakaszok arányát változtathatatlanak (invariánsnak) tekintjük, akkor épkezláb matematikai modellt nyerünk, amely kiemeli az euklideszi tér metrikától független tulajdonságait. Ez az affin geometria. Ha a párhuzamosság tulajdonságától is elvonatkoztatunk, de megőrizzük a „lineáris alakzatnak lenni”, az illeszkedés és a kettőviszony tulajdonságait még absztraktabb matematikai modellhez, a projektív geometriához jutunk. Ha az euklideszi, nem-euklideszi és a különböző differenciálgeometriai terek közös tulajdonságait vizsgáljuk (eközben a terek mindig mint konkrét, létező objektumok jelennek meg tudatunkban) eljutunk oda, hogy e terekben vannak nyílt és zárt halmazok, illetve a pontnak környezetei. Így keletkezik az e terek közös tulajdonságait tükröző újabb absztrakció, a topologikus tér fogalma.

Bizonyos elektromos jelfogókból felépített áramkörök matematikai modelljeként megjelenik az ún. „kapcsolásalgebra”. Az emberi gondolkodás bizonyos elemi jellegzetességeit tükrözi matematikailag az ítéletkalkulus. A matematika alapjainak tanulmányozása során megszületik a halmazelmélet és ennek elemi részeként a halmazalgebra. A „kapcsolásalgebra”, az ítéletkalkulus és a halmazalgebra matematikai modellek. Létrejöttük után azonban az objektív valóság részeivé válnak és mint ilyenek tanulmányozhatók. Mindhárom modell közös tulajdonságait tükrözi az absztrakció következő lépcsőjén a *Boole-algebra*.

Hasonló absztrakciós lánc a következő: a heliocentrikus rendszer — *Kepler törvényei* — $mr'' = -\gamma Mm/r^2$ — a differenciálegyenletek elmélete.

3. A matematika tárgya és helye a tudományok rendszerében

A tudományok feladata és célja a valóságban lejátszódó jelenségek, folyamatok megismerése, megmagyarázása, az objektív törvények feltárása és közvetlenül, vagy közvetve az emberiség szolgálatába állítása. Minden tudomány tárgya az objektív valóság és az előző fejezetben láttuk, hogy ez alól a matematika sem kivétel.

A valóság azonban bonyolult, hiszen „minden mindennel összefügg”, a legegyszerűbb valóságos jelenségek is összetettek. Az emberi tudat nem képes a valóságot, a valóságos jelenségeket a maguk teljes összetettségében, egyszerre visszatükrözni. Ezen a nehézségen a tudomány úgy segít, hogy egyrészt absztrahál, elvonatkoztat, másrészt a valóságot szeletekre vágja és az egyes szeleteket külön-külön tanulmányozza. Így differenciálódott az ókorban még egységes természettudomány matematikára, csillagászatra, mechanikára, fizikára, kémiára, biológiára, geológiára, műszaki tudományokra stb. Hasonló differenciálódás történt a társadalomtudományokban is. Ez az elkülönülési folyamat napjainkban is tart. Jó geométer lehet

valaki akkor is, ha nem ismeri a funkcionálanalizist, jó asztrofizikus akkor is, ha nem ért az égi navigációhoz.

Ma a tudományokat általában tárgyak szerint határozzuk meg, aszerint, hogy a valóság melyik szeletét, milyen vonatkozását, melyik anyagi mozgásformát tanulmányozzák. Viszonylag egyszerű megmondani, hogy mi a tárgya például a csillagászatnak, a geológiának, a biológiának, a történettudománynak, a jogtudománynak, a szociológiának stb. Azokat a tudományokat, amelyek a természet jelenségeivel, a természetben lejátszódó folyamatokkal foglalkoznak, *természettudományoknak*, azokat, amelyek az ember által létrehozott anyagi valósággal, a termelés során szerzett tapasztalatok általánosításával foglalkoznak, *műszaki tudományoknak*, az emberi társadalom jelenségeivel foglalkozókat *társadalomtudományoknak* nevezzük. Különös szerepe van a *filozófiának*, mint a „legáltalánosabb törvényszerűségek” tudományának, amellyel itt nem kívánok foglalkozni.

Mi azonban a matematika tárgya, az anyag, mely mozgásformájával foglalkozik ez a tudomány, hol helyezhetjük el a tudományok rendszerében?

Ha az anyag különböző mozgásformáit és a velük foglalkozó tudományokat el akarjuk határolni egymástól, az objektív valóságot, a világot egy térképpel ábrázolhatjuk, amelyen a különböző mozgásformáknak (vagy a megfelelő tudományoknak) különböző „országok” felelnek meg. Bár ezen a térképen a határok helyenként elmosódottak, mivel határterületi tudományok szép számmal vannak és jelentőségük fokozódik, az „országok” elhatárolhatók egymástól. (A biológiát meg tudjuk különböztetni a kémiaától, bár biokémia is létezik.) Egy ilyen térképen a matematikát nem tudjuk külön „orszáként” ábrázolni. A matematikát ábrázoló halmaz átmetszi az összes, vagy legalábbis a legtöbb „országot”. A matematikának közös része van a mechanikával, a fizikával, a kémiával, a biológiával, a közgazdaságtannal, a szociológiával, a műszaki tudományokkal, a hadtudományokkal, a logikával, a történettudománnyal és számos más tudománnyal. A matematika ugyanis foglalkozik mindezen tudományok azon vonatkozásaival, amelyeknek kvantitatív tulajdonságaik, formáik vannak, sőt továbbmenve, mindazon tulajdonságokkal (kvalitatív tulajdonságokkal is) és jelenségekkel, amelyek valamilyen módon formalizálhatók. Lássunk néhány példát a sok közül az előbbieik alátámasztására és szorítkozzunk csupán 20. századi példákra.

A *Ljapunov-féle direkt módszer*, amely most másodvirágzását éli, mechanikai fizikai és műszaki rendszerek vizsgálatával kapcsolatban keletkezett, mostani fejlődését a szabályozási rendszereknek az utóbbi évtizedekben megnövekedett jelentősége motiválja, és eredményei közvetlenül alkalmasak műszaki stabilitási feladatok megoldására.

A *matematikai* (lineáris, dinamikus, sztochasztikus) *programozás* problémaköre elsősorban hadtudományi és közgazdasági feladatokból nőtt ki és eredményei egyben közgazdaságtudományi eredmények is.

Az *információelmélet* átviteltechnikai és kódolási feladatok megoldására született.

A valószínűségelmélet *bolyongási problémái* elsősorban a fizikai *Brown-mozgás* leírására keletkeztek.

A *differenciális játékok elmélete* a légvédelmi rakétarendszerek kidolgozásával kapcsolatban született.

A *populáció-dinamika* matematikai modelljei kezdve a *Lotka—Volterra—Kolmogorov-féle modellen* a matematika olyan újabb ágát alkotják, amelynek eredményei

közvetlen biológiai mondanivalót hordoznak. Így többek között a „versengő kizárás elvére” (competitive exclusion principle, amely szerint két, egyetlen közös erőforráson, táplálékon élő faj közül a kisebb életképességű kipusztul) új matematikai elmélet épült.

A *matematikai logika* eredetileg az emberi gondolkodás folyamatának formalizált, absztrakt matematikai modellje, de szerepe jelentős a számítástudományban is.

A példákat tovább lehetne sorolni. Ha tehát a tudományokat tárgyak szerint határozzuk meg, akkor a matematika meghatározása nehezebb, mint bármely más tudományé. *A matematika ugyanis a természet, a társadalom és a megismerés mindazon jelenségével, tulajdonságával, vonatkozásával foglalkozik, amely formalizálható. Ezért a matematika természettudomány is, társadalomtudomány is és gnoszeológia is, de nem sorolható be egyszerűen pl. a természettudományok közé.*

Ennek az igazságnak felismerését az is nehezíti, hogy amint ezt az előző fejezetben láttuk, a matematika a már megalkotott fejezeit besorolja az objektíve létező kategóriák közé és vizsgálat tárgyává teszi azokat.

Jogos a kérdés, hogy ha a matematika ennyi mindenfélével képes foglalkozni, vagyis ha tárgya gyakorlatilag szinte akármilyen lehet, akkor hogyan lehet megmondani azt, mi matematika és mi nem az. Ez a kérdés gyakorlati szempontból is fontos, ha adott esetben el akarjuk dönteni egy kutatási eredményről, hogy az matematikai eredmény-e.

Ezzel a kérdéssel kapcsolatban foglalkoznunk kell a *matematika módszerével*. Egy kutatási eredményt ugyanis a követett módszer avat matematikai eredménnyé és nem az, hogy a kutatás tárgya természeti, társadalmi, a megismerés folyamatával kapcsolatos, vagy (ez utóbbi esetben) esetleg éppen matematikai objektum-e. A módszer lényege pedig a következő. A vizsgált objektum, jelenség, folyamat jellemző és számunkra érdekes kvantitatív, vagy kvalitatív tulajdonságait *szimbólumokkal* ábrázoljuk. A szimbólumok egy részét esetleg *primitív fogalomként* vesszük fel és az utóbbiak közötti alapvető összefüggéseket *axiómák* formájában mondjuk ki. Az axiómáknak egyértelműeknek és egymással ellentmondásban nem állóknak kell lenniük. Nem szerepelhet bennük más, mint primitív, vagy már *definiált fogalom*. A többi szimbólumot a primitív fogalmak, az axiómák és a már definiált fogalmak segítségével pontosan defináljuk. A továbbiakban e szimbólumokat eredetükről mintegy megfeledkezve úgy kell kezelniük, hogy csupán az axiómákban és definíciókban kimondott tulajdonságaikat és összefüggéseiket használjuk fel. Ezek után minden további összefüggést egzakt logikai úton, deduktív módon *bizonyítanunk* kell. Az ily módon elért eredmény matematikai, vagy matematikai *is*. Természetesen lehetséges, hogy az út egy részét már bejárták a kutató előtt, és az objektum tulajdonságait jelképező szimbólumok és a közöttük levő alapvető összefüggések beleillenek egy már korábban megalkotott matematikai diszciplína kereteibe. Ebben az esetben „csupán” újabb összefüggések feltárása, újabb tételek bebizonyítása a feladat.

Az előző bekezdésben azt mondtuk, hogy ha az elmélet keretei, az axiómarendszer már megvan, akkor minden további összefüggést, tételt *deduktív* úton bizonyítani kell. Sokan ezt félreértik és azt mondják, hogy a mai matematika tisztán deduktív tudomány.

Tisztán deduktív tudomány nincs, és a matematika sem az. Ezt mindenki tapasztalta, aki valaha is matematikai kutatást folytatott, csak talán tapasztalata nem jutott a tudatáig. Ahhoz hasonlóan, ahogyan MOLIÈRE úrhatnám polgára sem tudta,

míg erről föl nem világosították, hogy világéletében *prózában* beszélt. Az induktív okoskodási módszer a matematikai kutatásban elsősorban a következő két fázisban és formában van jelen.

Először az elvonatkoztatás, az absztrakció folyamatában, új matematikai elmélet megalkotása esetében, az axiómarendszer összeállításában. Ilyenkor a matematikus a vizsgált objektum tulajdonságaival való foglalkozás során szerzett tapasztalat, kísérletezés és intuíció révén választja el azokat a tulajdonságokat, amelyeket figyelembe kell venni, azoktól, amelyektől el kell vonatkoztatni. BOLYAI és LOBACHEVSKIJ az elődök és saját sikertelen *kísérleteik* után, melyek során bizonyítani igyekeztek a párhuzamosok axiómáját a többi euklideszi axióma alapján, e negatív tapasztalatok birtokában jutottak arra a következtetésre, hogy ez az axióma független a többitől. Induktív módszer vezetett tehát oda, hogy a párhuzamosok euklideszi axiómájának ellentettjét vegyék fel az axiómarendszerbe és megnézzék, valóban ellentmondásmentes rendszert kapnak-e. LJAPUNOV a mechanikai rendszerekkel kapcsolatos sok évszázados *tapasztalatot*, hogy ti. ott van a rendszernek stabilis egyensúlyi helyzete, ahol az energiának minimuma van, használta fel és általánosította akkor, amikor a *definit Ljapunov-függvények* bevezetésével kritériumokat adott a stabilitásra olyan esetekben is, amikor energiáról, sőt mechanikáról szó sem volt.

Másodszor új tételek kimondásában és bizonyításában. Hacsak nem egészen triviális állításról van szó, a matematikus gyakran példákon, speciális eseteken keresztül próbálja megsejteni, milyen jellegű állítás lehet igaz és csak ezután kísérli meg az állítás igazolását a lehetőség szerint legkevesebb feltétel felhasználásával. Ennek a tipikus matematikai kutatási módszernek tömör kifejezése a jelszó: „sejts és bizonyíts!” *Itt a sejtés a munka induktív fázisának, a bizonyítás a deduktív fázisnak felel meg.* (Természetesen előfordul, hogy másképpen kell dolgozni; éppen a probléma általánosítása hozza meg az eredményt.)

PÓLYA GYÖRGY erről a kérdésről a következőket írja: „A matematikát deduktív tudománynak tekintik. Ez azonban a matematikának csak az egyik oldala. A kész matematika végleges formájában tisztán deduktívnek tűnik, amely csak bizonyításokból áll. A matematika az alkotás folyamatában azonban hasonlít bármely más emberi ismeret megszerzésének folyamatára. Meg kell sejtened a matematikai tételt; mielőtt bebizonyítod; meg kell sejtened a bizonyítás alapötletét, mielőtt kidolgoznád a részleteket. Megfigyeléseket kell összehasonlítani, analógiákat kell követned; újból és újból próbálkoznod kell. A matematikus alkotó munkájának eredménye egy deduktív gondolatmenet, egy bizonyítás; azonban a bizonyításhoz plauzibilis okoskodás, találgatás vezet. ...Észre kell vennünk azt (amit a modern szerzők szinte elfelejtettek, de a régiek, mint EULER és LAPLACE világosan láttak), hogy az induktív okoskodás ahhoz hasonló szerepet játszik a matematikai vizsgálatokban, mint a fizikai kutatásban.” (*Induction and Analogy in Mathematics*, Princeton U.P. 1954.)

4. Alkalmazott és tiszta matematika

E dolgozat témájával kapcsolatban foglalkoznunk kell a tiszta és az alkalmazott matematika kérdésével egyrészt azért, mert a matematika kapcsolatát a matematikán kívüli objektív valósággal elsősorban az alkalmazott matematika teremti meg, másrészt azért, mert ebben a kérdésben számos téves nézet lát napvilágot. Egyesek tagadják azt, hogy a matematikát egyáltalán fel lehet osztani tiszta és alkalmazott

részre, mások elismerik azt, hogy létezik alkalmazott matematika, de azt nem tekintik matematikának.

A kiigazodást nehezíti az, hogy országonként, hagyományosan más-más értelmet tulajdonítanak az alkalmazott matematikának mint tantárgynak a felsőoktatási intézményekben. A brit egyetemeken az applied mathematics magában foglalja a matematikai statisztikát, a mechanikát és azokat a tudományterületeket, amelyeket nálunk az elméleti fizikához szokás sorolni. A szovjet egyetemeken a прикладная математика elsősorban a numerikus analízist és a számítástudományt jelenti.

Ha eltekintünk a nemzeti tradícióktól és a dolgot kizárólag a kutatás, illetve a kutatási eredmények szemszögéből nézzük, úgy tűnik, megvalósítható a matematikai kutatások és eredmények felosztása alkalmazott, illetve tiszta matematikai kutatásokra és eredményekre. Le kell azonban előljáróban szögezni, hogy minden ilyen felosztás relatív és bizonyos mértékig szubjektív. Relatív azért, mert egy eredmény, amelyet ma a tiszta matematikához sorolunk, egy későbbi időpontban alkalmazott matematikáinak bizonyulhat és fordítva. Szubjektív azért, mert egy kutatómunka, vagy tudományos eredmény ide-, illetve odasorolásában szerepet játszik a kutatást végzők, vagy az eredményt elérők szándéka is.

Az előző fejezetben kifejtettük, hogy egy tudományos eredményt akkor tekintünk matematikainak, ha *módszerében* matematikai, függetlenül tárgyától. A matematikai eredményeket aszerint soroljuk a tiszta, illetve az alkalmazott kategóriába, hogy milyen *tárgyra*, milyen célra irányult a kutatás.

Alkalmazott matematikai kutatásnak, illetve eredménynek nevezünk egy tudományos munkát, ha a kutatás tárgya a matematikán kívüli objektív valóság és a módszere matematikai. Tiszta matematikai kutatásnak, illetve eredménynek nevezünk egy tudományos munkát, ha a kutatás tárgya a matematikán belüli valóság és a módszere matematikai.

Ha tehát a kutatás módszere matematikai és célja pl. egy fizikai modell, egy biológiai vagy műszaki rendszer, egy közgazdasági probléma matematikai modelljének megalkotása, vagy egy ilyen modellnek az analízise, vagy továbbfejlesztése, vagy az így nyert eredmények összevetése a gyakorlati, kísérleti tapasztalattal, akkor alkalmazott matematikai tudományos kutatásról beszélünk. Ha viszont a kutatás célja egy a matematika valamely már kidolgozott fejezetében felerült probléma megoldása, vagy éppen egy matematikai elmélet általánosítása, vagy egy matematikai eredmény alkalmazása a matematika egy másik fejezetében anélkül, hogy ezt a célt a matematikán kívüli objektív valóság megismerése konkrét formában és közvetlenül motiválná, akkor tiszta matematikai kutatásról beszélünk.

Félreértések elkerülése végett meg kell különböztetnünk az alkalmazott matematikusi tevékenységet az alkalmazott matematikai tudományos eredménytől. Előfordulhat, hogy az alkalmazott matematikusi tevékenység, az alkalmazott matematikai kutatómunka egy másik tudomány szempontjából tudományos eredményre vezet, de ez nem minősíthető alkalmazott matematikai tudományos eredménynek. Akkor következik ez be, ha a vizsgált problémát olyan kész, kidolgozott matematikai apparátussal lehet modellezni, amely nem szorul továbbfejlesztésre. Más szóval akkor, ha a feladat megoldásához nem kell a matematika módszerét alkalmaznunk, mivel ezt már mások elvégezték előttünk. Ilyenkor magának az alkalmas modellnek a megtalálása (pl. az irodalomban) jelentős lehet tudományos szempontból, de nem alkalmazott matematikai eredmény.

Amint erre korábban utaltunk, a matematika története folyamán gyakran megtörtént, hogy egy tiszta matematikai eredmény később kész alkalmazott matematikai modellként szolgált. Ilyen példák szerepeltek a 2. fejezetben. Még gyakrabban fordul elő, sőt az a tipikus, hogy alkalmazott matematikai eredményként született elmélet tiszta matematikai kutatások kiinduló pontjává és inspirálójává válik. A számos ilyen jellegű példa közül egyik legjelentősebb maga a differenciál- és integrálszámítás, amely Newtonnál kifejezetten mechanikai köntösben jelentkezett. Egy másik jelentős példa az, ahogyan FOURIER a róla elnevezett sorokat bevezette *Théorie analytique de la chaleur* c. művében. Később a *Fourier-sor* az ortogonális sorok és a *Hilbert-terek* elméletének kiinduló pontjává vált. Tévedés lenne azt hinni, hogy ilyen jellegű példák a 20. században nem lennének, vagy akár csak gyakoriságuk és jelentőségük csökkenne. Elegendő talán csak a valószínűségelméleti módszerek számelméleti alkalmazásaira, vagy a topologikus dinamika kialakulására utalni.

A tiszta, illetve az alkalmazott matematika meghatározásáról az előbbiekben röviden vázolt elképzelésemet bizonyára javítani, pontosítani lehet. A tiszta és az alkalmazott matematika értelmezését, egymáshoz való viszonyát, szerepét azonban végre tisztáznunk kell, mivel a matematika e két összetevőjének harmonikus kiegyensúlyozott fejlődése nélkülözhetetlen a tudomány egészének egészséges előrehaladásához.

Ha egy ország tudományos életében, vagy akár egy nagyobb tudományos kutatóhelyen csak az alkalmazott matematikának van jelentősége, és a tiszta matematikai kutatások csenevészek, jelentéktelenek, nem eléggé széleskörűek, akkor egy önmagát erősítő, negatív tendencia válik uralkodóvá. Egyre inkább szűk, prakticista, az elmélet jelentőségét tagadó szemléletmód terjed el, új tehetségek felfedezése és kibontakozása lelassul. Maguknak az alkalmazott matematikai kutatásoknak a színvonala is esik az általánosítás igényének hiánya miatt. Vannak olyan nagy tudományos múlttal rendelkező országok, amelyek riasztó például szolgálhatnak a fentiekre.

Hasonlóan súlyos következményre vezet az, ha egy ország, vagy egy nagyobb intézmény tudományos életében túlnyomó részt csak a tiszta matematika van jelen, az alkalmazott matematikai kutatások gyengék és nem megbecsültek. Először is ez a helyzet lefékezi „a tudomány termelő erővé válásának” folyamatát, és így a társadalom nem látja közvetlen hasznát és értelmét a matematikai kutatásoknak. Ennek következtében csökken a matematika tudományának a társadalomban betöltött szerepe, jelentősége. A gazdasági életben ott is ad hoc empirikus eljárásokat alkalmaznak, ahol matematikai módszereket is lehetne. Ott ahol, a matematikai módszerek alkalmazása mégis elkerülhetetlen, ez gyakran kevésbé képzett, gyengébb képességű személyek, néha sarlatánok kezébe kerül. Másodszor, ha a tiszta matematika nem kap rendszeresen ösztönzést az alkalmazott matematikai kutatásoktól, nem kap olyan reális problémákat, amelyek a valóság tanulmányozása során vetődnek csak fel, akkor fejlődése lelassul, lényeges központi jelentőségű problémák helyett, periférikus témák kerülnek előtérbe, a matematika „elbarokkosodik”.

(Beérkezett: 1982. július 6.)

FARKAS MIKLÓS

BME GÉPÉSZMÉRNÖKI KAR MATEMATIKA TANSZÉK
1521 BUDAPEST, STOCZEK U. H. ÉP. IV. E. 42.

A FÜGGŐSÉGEK ÉS AZ INDIVIDUUMOK SZÁMA KÖZÖTTI KAPCSOLAT ÖSSZETETT ADATRENDSZEREK ESETÉN

DEMETROVICS JÁNOS

FÜREDI ZOLTÁN

KATONA GYULA

Budapest

Egy korábbi dolgozatunkban azt a feladatot vizsgáltuk, hogy egy adott kulcsrendszerrel rendelkező adatbázisban mennyi az individuumok minimális száma. A jelen dolgozatban részben ezzel a problémával, részben egy természetes analogonnal foglalkozunk. Azt is vizsgálni fogjuk ugyanis, hogy mennyi az individuumok minimális száma, ha az adatbázisban a függőségek rendszere is adott.

Sokszor hasznos az, ha az összetett adatbázis bizonyos paramétereit ki tudjuk számítani, vagy meg tudjuk becsülni a rész-adatbázisok megfelelő paramétereiből. Ezt kíséreljük meg elvégezni az említett paraméterre, az individuumok minimális számára. Erre jelenleg csak akkor vagyunk képesek, ha a nagy adatbázis egymástól független kisebbekből van felépítve.

1. Bevezetés

Az adatbázis struktúrájának a felhasználó számára legfontosabb jellemzője az, hogy a tárolt adatok közötti összefüggések milyen formában vannak jelen az adatbázisban. Mivel egy adatbázis struktúráját a rekordtípusok és a rekordtípusok közötti kapcsolatok alkotják, ezért az adatmodelleket általában aszerint osztályozzuk, hogy a rekordtípusok között milyen kapcsolat definiálása megengedett. Ennek alapján lényegében két adatmodellt különböztetünk meg:

- az egyik az ún. *hálózatos adatmodell*,
- a másik az ún. *relációs adatmodell*.

Rekordtípus vagy más néven adatbázis megadása matematikai szempontból nem más, mint egy „értelmezési tartomány” és egy „értékkészlet” definíálása. Az így kapott rekordtípus előfordulásai az adott értelmezési tartományon értelmezett, értékeit az adott értékkészletből felvevő függvények. Az értelmezési tartomány elemeit attributumoknak nevezik. Egy adatbázis egy rekordtípusának egy adott pillanatban létező előfordulásai tehát egy mátrixot alkotnak. Ennek sorai az egyes előfordulások, oszlopai pedig az egyes attributumokon az előfordulások értékei.

Egy korábbi dolgozatunkban [2] vizsgáltuk azt a feladatot, hogy egy adott kulcsrendszerrel rendelkező adatbázisban mennyi az individuumok minimális száma. A jelen dolgozatban részben ezzel a problémával, részben egy természetes analogonnal foglalkozunk. Azt is vizsgálni fogjuk ugyanis, hogy mennyi az individuumok minimális száma, ha az adatbázisban a függőségek rendszere is adott.

Egy adatbázis sokszor természetes módon épül fel több adatbázisból. Máskor egy adatbázist mesterséges módon igyekszünk felbontani kisebb adatbázisokra. Mindkét esetben hasznos, ha az összetett adatbázis bizonyos paramétereit ki tudjuk számítani, vagy meg tudjuk becsülni a rész-adatbázisok megfelelő paramétereiből. Ezt kíséreljük meg elvégezni az említett paraméterre, az individuumok minimális

számára. A kutatás jelenlegi állásában erre is csak akkor vagyunk képesek, ha a nagy adatbázis egymástól független kisebbekből van felépítve. Természetesen ugyanerre szükség volna azon esetben is, amikor az összeépítés bonyolultabb.

A 4. fejezetben a fentiekkel kapcsolatos további gondolatok vannak.

2. Definíciók, alapfogalmak

Az adatbázis egyik legfontosabb fogalma az E. F. CODD [3] által definiált funkcionális függés. Az előzőekből megállapíthatjuk, hogy egy relációs adatbázis pillanatnyi tartalmát mátrixnak tekinthetjük, ezért a funkcionális függést mátrixokra definiáljuk.

Lássuk most a pontos matematikai definíciókat. Egy *adatbázist* úgy tekintünk, mint egy m sorból és n oszlopból álló mátrixot, amelynek értékei nemnegatív egész számok, és amelynek minden sora különböző. A sorok felelnek meg az indiduumoknak, az oszlopok az attributumoknak. Az oszlopok egy halmaza *kulcs*, ha nincs két olyan különböző sor, amelyek megegyeznek ezekben az oszlopokban. Vagyis a kulcs oszlopaiban álló értékek egyértelműen meghatározzák a sort. Ennél általánosabban, ha A és B az oszlopok egy-egy halmaza, és minden két olyan sor, amely megegyezik A -ban, az megegyezik B -ben is, akkor azt mondjuk, hogy A -ból *következik* B (jelölésben $A \rightarrow B$). Az ilyen (A, B) párokat függőségeknek nevezzük. Jól látható, hogy ha A egy kulcs és X az összes oszlopok halmaza, akkor $A \rightarrow X$. Egy A kulcsot *minimális kulcsnak* nevezünk, ha nincs olyan B kulcs, melyre $B \subset A$, $B \neq A$ teljesülne. A minimális kulcsok rendszerét \mathcal{K} -val jelöljük. Ismeretes [4], hogy \mathcal{K} akkor és csak akkor lehet valamely adatbázis kulcsainak rendszere, ha nem üres *Sperner-rendszer*, azaz nincs két különböző A, B eleme, amelyekre $A \subset B$ fennállna.

Legyen az M adatbázishoz hozzárendelve a következő X részhalmazából Y részhalmazába leképező függvény:

$$(2.1) \quad \mathcal{L}_M(A) = \{b: A \rightarrow \{b\}\}$$

Könnyen látható, hogy ez az \mathcal{L}_M függvény egyrészt egyértelműen meghatározza a függőségek teljes rendszerét ($A \rightarrow B \Leftrightarrow B \subseteq \mathcal{L}_M(A)$) másrészt rendelkezik a következő tulajdonságokkal:

$$(2.2) \quad A \subseteq \mathcal{L}(A)$$

$$(2.3) \quad A \subseteq B \Rightarrow \mathcal{L}(A) \subseteq \mathcal{L}(B),$$

$$(2.3a) \quad \mathcal{L}(\mathcal{L}(A)) = \mathcal{L}(A),$$

a fenti három tulajdonsággal rendelkező \mathcal{L} függvényeket *lezárásoknak* nevezzük. Ismeretes, hogy minden \mathcal{L} lezáráshoz van olyan adatbázis, amelyben a függőségek által generált \mathcal{L}_M függvény éppen \mathcal{L} -vel egyenlő.

Legyen \mathcal{K} az n -elemű halmaz részhalmazainak egy *Sperner-rendszere*. Mint fent említettük, létezik olyan adatbázis, amelyben ez a minimális kulcsok rendszere. Legyen ezen adatbázisok közül a minimális sorból állónak $s(\mathcal{K})$ sora. [1]-ben található $s(\mathcal{K})$ -ra vonatkozó eredmények néhány egyszerű \mathcal{K} rendszer esetén. A fentiekkel analóg módon definiálhatjuk $s(\mathcal{L})$ -et: Egy adott \mathcal{L} lezáráshoz található adatbázisok közül a minimális sorúak sorszámaként.

Ha \mathcal{L}_1 egy n_1 elemű X_1 halmazon egy lezárás, \mathcal{L}_2 pedig egy n_2 elemű X_2 -n ($X_1 \cap X_2 = \emptyset$), akkor definiáljuk az $n_1 + n_2$ elemű $X_1 \cup X_2$ -n az \mathcal{L}_1 és \mathcal{L}_2 lezárások $\mathcal{L}_1 \times \mathcal{L}_2$ direkt szorzatát a következőképpen:

$$(2.4) \quad (\mathcal{L}_1 \times \mathcal{L}_2)(A) = \mathcal{L}_1(A \cap X_1) \cup \mathcal{L}_2(A \cap X_2).$$

Hasonlóan, ha X_1 -en, illetve X_2 -n csak a minimális kulcsok \mathcal{K}_1 , illetve \mathcal{K}_2 rendszere adott, akkor $\mathcal{K}_1 \times \mathcal{K}_2$ jelenti $X_1 \cup X_2$ azon részalmazainak rendszerét, amelyek egy \mathcal{K}_1 -beli és egy \mathcal{K}_2 -beli uniójaként állnak elő.

Ezek után kimondhatjuk főeredményünket:

2.1. TÉTEL.

$$s(\mathcal{L}_1 \times \mathcal{L}_2) = s(\mathcal{L}_1) + s(\mathcal{L}_2) - 1.$$

Bizonyítás. 1. Bizonyítsuk először a \leq egyenlőtlenséget. Legyen \mathbf{M}_1 egy olyan $s(\mathcal{L}_1)$ -szer n_1 -es mátrix, amelyre $\mathcal{L}_{\mathbf{M}_1} = \mathcal{L}_1$, azaz amely \mathcal{L}_1 -et a legkevesebb sorral realizálja. Hasonlóan, legyen \mathbf{M}_2 az \mathcal{L}_2 lezárást legkevesebb sorral realizáló mátrix. Ezek segítségével fogunk egy $s(\mathcal{L}_1) + s(\mathcal{L}_2) - 1$ sorú és $n_1 + n_2$ oszlopú \mathbf{M} mátrixot konstruálni, amelyre $\mathcal{L}_{\mathbf{M}} = \mathcal{L}_1 \times \mathcal{L}_2$. \mathbf{M} konstrukcióját a következő séma szemlélteti:

\mathbf{M}_1	β
	\vdots
	β
	β
	β
α	\mathbf{M}_2
α	
α	
\vdots	
α	

1. ábra

Azaz \mathbf{M}_1 -et és \mathbf{M}_2 -t úgy helyezzük egymás mellé, hogy csak egy sorban találkozzanak, \mathbf{M}_1 utolsó sora, α legyen \mathbf{M}_2 első sorával, β -val egy vonalban. Ezután \mathbf{M}_2 többi sora mellé α -t, \mathbf{M}_1 többi sora mellé β -t írjunk. Triviális, hogy \mathbf{M} -nek $s(\mathcal{L}_1) + s(\mathcal{L}_2) - 1$ sora és $n_1 + n_2$ oszlopa van, tehát csak azt kell bizonyítani, hogy az \mathbf{M} -hez tartozó $\mathcal{L}_{\mathbf{M}}$ lezárás $\mathcal{L}_1 \times \mathcal{L}_2$. Vagyis, ha $A \subset X = X_1 \cup X_2$, $a \in X$, akkor

$$(2.5) \quad a \in \mathcal{L}_{\mathbf{M}}(A) \Leftrightarrow a \in \mathcal{L}_1(A \cap X_1) \cup \mathcal{L}_2(A \cap X_2).$$

A szimmetria miatt feltehetjük, hogy $a \in X_1$. Ekkor (25) úgy is fogalmazható, hogy a

$$(2.6a) \quad a \in \mathcal{L}_1(A \cap X_1) \Rightarrow \mathbf{M} \text{ minden olyan két sora, amely } A\text{-ban megegyezik, megegyezik } a\text{-ban is}$$

és

$$(2.6b) \quad a \notin \mathcal{L}_1(A \cap X_1) \Rightarrow \mathbf{M}\text{-nek van két olyan sora, melyek } A\text{-ban megegyeznek, de } a\text{-ban nem}$$

feltételeknek együtt kell teljesülnie. Bizonyítsuk be tehát ezeket. Ha $a \in \mathcal{L}_1(A \cap X_1)$, ekkor \mathbf{M}_1 definíciója szerint \mathbf{M}_1 minden olyan két sora, amely $A \cap X_1$ -en meg-

egyezik, a -n is megegyezik. Ha \mathbf{M} kiválasztott két sora (ami A -ban megegyezik) mindkettő α -val kezdődik, nincs mit bizonyítanunk. Ha egyik se kezdődik α -val, akkor \mathbf{M}_1 fenti tulajdonságából következik az a -ban való megegyezés. Ha egyik α -val kezdődik, a másik nem, akkor ismét \mathbf{M}_1 fenti tulajdonsága elegendő, hiszen α az \mathbf{M}_1 egy sora.

Térjünk most rá (2.6b) bizonyítására. Tegyük fel, hogy $a \notin \mathcal{L}_1(A \cap X_1)$. \mathbf{M}_1 -nek definíció szerint létezik két olyan sora, melyek $A \cap X_1$ -ben egyeznek, de a -ban nem. Az ezeknek megfelelő két sor \mathbf{M} -ben kielégíti (2.6b) jobb oldalát.

2. $A \cong$ egyenlőtlenség bizonyításához vegyünk egy tetszőleges \mathbf{M} mátrixot, melyre $\mathcal{L}_{\mathbf{M}} = \mathcal{L}_1 \times \mathcal{L}_2$ és bizonyítsuk be, hogy \mathbf{M} sorainak száma $\cong s(\mathcal{L}_1) + s(\mathcal{L}_2) - 1$. Ennek bizonyításához két lemmára lesz szükségünk. Az \mathbf{M} mátrix \mathcal{L}_1 -nek megfelelő első n_1 oszlopát jelöljük \mathbf{M}_1 -gyel, a maradékot \mathbf{M}_2 -vel.

2.1. LEMMA. $\mathcal{L}_{\mathbf{M}_2} = \mathcal{L}_2$.

Bizonyítás. Legyen $A \subset X_2$, $a \in X_2$. Lássuk be először, hogy ha $a \in \mathcal{L}_2(A)$, akkor $a \in \mathcal{L}_{\mathbf{M}_2}(A)$. A feltételből és \mathbf{M} definíciójából következik, hogy ha \mathbf{M} két sora megegyezik A -ban, akkor a -ban is. Ez természetesen az \mathbf{M}_2 részmátrixban is igaz. Megfordítva, ha $a \notin \mathcal{L}_2(A)$, akkor $a \notin (\mathcal{L}_1 \times \mathcal{L}_2)(A)$, így \mathbf{M} -nek van két olyan sora, amelyek A -ban megegyeznek, de a -ban nem. Ugyanez a két sor \mathbf{M}_2 -ben is megfelel, azaz $a \notin \mathcal{L}_{\mathbf{M}_2}(A)$.

Természetesen $\mathcal{L}_{\mathbf{M}_2} = \mathcal{L}_1$ ugyanígy igaz. Azonban \mathbf{M}_1 -re egy kissé erősebb állítást fogunk használni:

2.2. LEMMA. Legyen \mathbf{N} egy mátrix, $\mathcal{L}_{\mathbf{N}}$ az általa generált lezárás. Tegyük fel, hogy \mathbf{N} sorai k osztályba sorolhatók úgy, hogy ha $a \notin \mathcal{L}_{\mathbf{N}}(A)$, akkor található 2. egy osztályba tartozó sora \mathbf{N} -nek, melyek A -ban megegyeznek, de a -ban nem. Ekkor

$$(2.7) \quad \mathbf{N} \text{ sorainak száma } \cong s(\mathcal{L}_{\mathbf{N}}) + k - 1.$$

Bizonyítás. A bizonyítást k -ra vonatkozó indukcióval végezzük. Ha $k=1$, (2.7) definíció szerint igaz. Tegyük fel tehát, hogy k -ra igaz és bizonyítsuk be $k+1$ -re. \mathbf{N} sorai legyenek $k+1$, a lemma feltételeinek megfelelő osztályba sorolva. Ha \mathbf{N} -nek konstans oszlopai lennének, hagyjuk el ezeket. Ez módosított mátrix, amelyet továbbra is \mathbf{N} -nel jelölünk, a lemma feltételeit változatlanul teljesíti, $\mathcal{L}_{\mathbf{N}}(\emptyset) = \emptyset$. Változtassuk meg \mathbf{N} értékeit úgy, hogy két különböző osztályban egyáltalán ne legyen azonos érték, viszont egy osztályon belül az azonosság-különbözőség ne változzon (pl. az i -edik osztályban levő sor egy j eleme helyett írjunk $j \cdot p^i$ -t, ahol p egy olyan prímszám, ami minden \mathbf{N} -ben szereplő értéknél nagyobb). Az így keletkezett mátrixot jelöljük \mathbf{N}' -vel. Gondoljuk át, hogy

$$(2.8) \quad \mathcal{L}_{\mathbf{N}'} = \mathcal{L}_{\mathbf{N}}.$$

Tegyük fel, hogy $a \in \mathcal{L}_{\mathbf{N}}(A)$, azaz, ha \mathbf{N} két sora A -ban azonos, akkor a -ban is. Ekkor, ha \mathbf{N}' két sora azonos A -ban (akkor $A \neq \emptyset$ esetén \mathbf{N} azonos osztályából valók), akkor ugyancsak egyformák lesznek a -ban is. Így $a \in \mathcal{L}_{\mathbf{N}'}(A)$. Másrészt, ha $a \notin \mathcal{L}_{\mathbf{N}}(A)$, akkor \mathbf{N} -nek van két, egy osztályba eső sora, amelyek A -ban egyeznek, a -ban nem. Mivel egy osztályban vannak, ugyanez \mathbf{N}' -ben is teljesülni fog rájuk. Következésképpen $a \notin \mathcal{L}_{\mathbf{N}'}(A)$.

Legyen N' első és második osztályának egy-egy sora $\gamma = (\gamma_1, \dots, \gamma_u)$, illetve $\delta = (\delta_1, \dots, \delta_u)$. Hagyjuk el a mátrixból γ -t és írjunk az első osztály i -edik oszlopában szereplő minden γ_i helyett δ_i -t. Az így kapott N'' mátrixra igazoljuk most az

$$(2.9) \quad \mathcal{L}_{N''} = \mathcal{L}_{N'}$$

állítását. Tegyük fel, hogy $a \in \mathcal{L}_{N'}(A)$ és legyen N'' két sora A -ban azonos. Ha ez a két sor N' szerinti azonos osztályokban van, akkor a két sor a -ban való azonos-sága következik. Ha nem, akkor csak a volt első és második osztály egy-egy sora lehet, μ'' , illetve γ'' . Itt a két sor i -edik eleme csak akkor lehet azonos, ha mindkettő δ_i . De akkor a megfelelő N' -beli sorban, μ' -ben γ_i áll az i -edik helyen. Azaz, μ' az A -ba eső részében megegyezik γ -val. $a \in \mathcal{L}_{N'}(A)$ -ból következik, hogy e sor a -dik eleme N' -ben γ_a , N'' -ben δ_a . Hasonlóan látható, hogy a v' sor a -dik eleme δ_a , vagyis μ'' és v'' megegyezik az a -dik helyen. $a \in \mathcal{L}_{N''}(A)$.

Megfordítva, tegyük fel, hogy $a \notin \mathcal{L}_{N'}(A)$, azaz N' -nek létezik két sora, amelyek A -ban egyeznek, a -ban nem. Ezek természetesen azonos osztályba fognak esni. Ez a két sor változatlanul megfelel N'' -ben is. $a \notin \mathcal{L}_{N''}(A)$ következik. Ebből (2.9)-en kívül az is látható, hogy N'' -re a lemma feltevése már k -val teljesül, alkalmazható az idukciós feltevés:

$$(N \text{ sorainak száma}) - 1 = N'' \text{ sorainak száma} \cong s(\mathcal{L}_{N''}) + k - 1$$

A jobb oldalon (2.8) és (2.9) miatt $s(\mathcal{L}_N) + k - 1$ is írható, azaz

$$N \text{ sorainak száma} \cong s(\mathcal{L}_N) + k,$$

amint azt bizonyítani kívántuk.

Térjünk vissza a tétel, pontosabban a \cong egyenlőtlenség bizonyítására. Legyen tehát valamely M -re $\mathcal{L}_M = \mathcal{L}_1 \times \mathcal{L}_2$ és az \mathcal{L}_1 és \mathcal{L}_2 alaphalmazának megfelelő részmatrixok M_1 és M_2 . Az első lemma szerint $\mathcal{L}_{M_1} = \mathcal{L}_1$, de ennél valamivel több is igaz. Ha $a \notin \mathcal{L}_1(A)$ ($A \subset X_1$), akkor $a \notin \mathcal{L}_1(A) \cup X_2 = (\mathcal{L}_1 \times \mathcal{L}_2)(A \cup X_2)$, azaz M -nek van két olyan sora, melyek $A \cup X_2$ -ben megegyeznek, a -ban nem. Azaz, M_1 -ben két olyan sor található, 1. melyekhez tartozó M_2 -beli rész megegyezik, 2. A -ban megegyeznek, 3. a -ban nem egyeznek meg. Vagyis M_1 sorait (M_2 különböző sorainak száma) számú olyan osztályba soroltuk, melyek kielégítik a második lemma feltételeit. Így (2.7) alapján

$$M_1 \text{ sorainak száma} \cong s(\mathcal{L}_1) + (M_2 \text{ különböző sorainak száma}) - 1.$$

Az első lemma alapján M_2 különböző sorainak száma $\cong s(\mathcal{L}_2)$. Innen a kívánt egyenlőtlenség következik.

3. Kulcsrendszerek direkt szorzata

Ha \mathcal{K}_1 , illetve \mathcal{K}_2 a minimális kulcsok rendszere a diszjunkt X_1 , illetve X_2 halmazokon, akkor $\mathcal{K}_1 \times \mathcal{K}_2$ álljon $X_1 \cup X_2$ azon halmazaiból, amelyek $A \neq B$, $A \in \mathcal{K}_1$, $B \in \mathcal{K}_2$ alakban állnak elő. Könnyen látható, hogy $\mathcal{K}_1 \times \mathcal{K}_2$ egy *Sperner-rendszer*, azaz van olyan lezárás, amelyben ezek a minimális kulcsok.

3.1. TÉTEL.

$$(3.1) \quad s(\mathcal{K}_1 \times \mathcal{K}_2) \cong s(\mathcal{K}_1) + s(\mathcal{K}_2) - 1.$$

Bizonyítás. Az egyenlőtlenség a 2.1. tétel bizonyításának első részével azonos módon bizonyítható. Az ott leírt konstrukció itt is megfelel. A tapasztalat azt mutatja, hogy (3.1)-ben csaknem mindig egyenlőtlenség áll, a 2.1. tétel esetéhez hasonlóan. Azonban a következő ellenpélda mutatja, hogy ez nem mindig igaz.

Legyen

$$X_1 = \{1, 2, 3, 4, 5\}, \quad X_2 = \{6, 7, 8, 9, 10\},$$

$$\mathcal{K}_1 = \{\{1, 2\}, \{3, 4\}, \{1, 5\}, \{2, 5\}, \{3, 5\}, \{4, 5\}\}$$

és \mathcal{K}_2 ezzel izomorf X_2 -n. A maximális nem kulcsok rendszere X_1 -en $\mathcal{K}^{-1} = \{\{5\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$.

Lássuk be először, hogy $s(\mathcal{K}_1) \geq 5$. Tegyük fel indirekte, hogy van egy 4 soros M_1 mátrix, amelynek kulcsrendszere \mathcal{K}_1 . Fel fogjuk használni a következő egyszerű lemmát:

3.1. LEMMA. Ha M kulcsrendszere \mathcal{K} , akkor \mathcal{K}^{-1} minden A eleméhez létezik egy sorpár M -ben, melyek pontosan A -ban egyeznek, és \mathcal{K}^{-1} különböző elemeihez különböző sorpárok tartoznak.

Megjegyzés: A lemma állítása impliciten megjelenik már [1]-ben is.

Bizonyítás. Mivel minden A nem-kulcshoz tartozik egy sorpár, melyek egyenlők A -ban, de nem teljesen egyenlők, így $A \in \mathcal{K}^{-1}$ esetén is van ilyen. Ha ez a két sor több helyen is egyenlő lenne nem csak A -ban, akkor A -nak volna egy olyan valódi bővítése, ami \mathcal{K}^{-1} -ben volna, ellentmondásban $A \in \mathcal{K}^{-1}$ -gyel.

Ebből máris következik, hogy $s(\mathcal{K}_1) \geq 4$, hiszen $\binom{3}{2} = 3 < 5 = |\mathcal{K}_1^{-1}|$. De ha M sorainak száma 4, akkor is egy kivételével minden sorpár egyezési halmaza valamelyik \mathcal{K}_1^{-1} -beli elem. Még egy egyszerű lemmát használunk:

3.2. LEMMA. Egy mátrix 3 sorának 3 egyezési halmazára A_1, A_2, A_3 -ra fennáll

$$(3.2) \quad (A_i \cap A_j) - A_k = \emptyset,$$

ahol i, j, k az 1, 2, 3 egy permutációja.

Bizonyítás. Triviális.

Az M 4 sorából képezhető 6 párból 5-höz tartozik \mathcal{K}_1^{-1} -beli halmaz. (Legyen a kivétel a (3, 4) sorpár.) Ez az 5 pár 2 darab sorhármast határoz meg. Ezekre teljesülni kell (3.2)-nek. Mivel \mathcal{K}_1^{-1} -ben nincs 3 halmaz, amelynek a metszete nem üres, így a 3 halmaznak páronként is diszjunktak kell lennie. 3 halmaz (szimmetriától eltekintve) csak egyféleképpen választható így ki: $\{1, 3\}, \{2, 4\}, \{5\}$. De ekkor a másik hármas egyértelmű: $\{1, 4\}, \{2, 3\}, \{5\}$, ugyanis a két hármasnak nem lehet egynél több közös halmaza. Szimmetriaokok miatt feltehető tehát, hogy az 1. és 2. sor megegyezési halmaza $\{5\}$, az 1. és 3.-é $\{1, 3\}$, a 2. és 3.-é $\{2, 4\}$, továbbá az 1. és 4.-é $\{2, 3\}$. De ekkor a 3. és 4. sor megegyezik $\{1, 2\}$ -ben, de $\{3, 4\}$ -ben nem, azaz \mathcal{K}_1^{-1} -nek kellene lennie egy $\{1, 2\}$ -t tartalmazó elemének. Ez az ellentmondás bizonyítja, hogy $s(\mathcal{K}_1) \geq 5$.

Ha (3.1)-ben egyenlőség állna, $s(\mathcal{K}_1 \times \mathcal{K}_1) = s(\mathcal{K}_1^2) = 9$ lenne. Egy konstrukcióval igazolni fogjuk, hogy $s(\mathcal{K}_1^2) \leq 8$.

Használjuk a következő jelöléseket:

$$\alpha = 0, 0, 0, 0, 0$$

$$\beta = 1, 1, 1, 1, 0$$

$$\gamma = 0, 1, 0, 1, 1$$

$$\delta = 1, 0, 0, 1, 1$$

A mátrix, amelyik $s(\mathcal{K}_1^2) \leq 8$ -at bizonyítja, a következő:

$$\mathbf{M} = \begin{matrix} & \alpha & \beta & \gamma & \delta \\ \alpha & \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \beta & \delta & \alpha \\ \gamma & \alpha & \beta & \delta & \alpha \\ \delta & \beta & \delta & \alpha & \beta \end{matrix}$$

Lássuk be, hogy \mathbf{M} minimális kulcsainak halmaza éppen \mathcal{K}_1^2 . Ehhez először azt kell igazolnunk, hogy a \mathcal{K}_1^2 -beli halmazok \mathbf{M} -ben kulcsok, azaz nincs két olyan sora \mathbf{M} -nek, melyek különbözők, de egy \mathcal{K}_1^2 -beliben vagy egy ilyennek bővítésében egyeznek meg. Vagyis két sor vagy bal vagy jobb oldalán nem \mathcal{K}_1 -beli halmazban vagy ilyennek bővítésében egyezik. Könnyen ellenőrizhető, hogy $\alpha, \beta, \gamma, \delta$ közül csak egy pár, γ és δ egyezik \mathcal{K}_1 -beliben vagy bővítésében. Természetesen az egyformák is ilyenek. A két oldal összehasonlítása mutatja, hogy valóban nincs két olyan teljes sor, mely mindkét oldalon \mathcal{K}_1 -beliben vagy bővítésében egyezne.

Másrészt azt is igazolni kell, hogy $(\mathcal{K}_1^2)^{-1}$ elemei nem kulcsok. Könnyen látható, hogy $(\mathcal{K}_1^2)^{-1}$ az olyan halmazokból áll, melyek egyik oldalon egy \mathcal{K}_1^1 -beliből, a másik oldalon a teljes alaphalmazból állnak. Az ilyen halmazokra kell találni sorpárt \mathbf{M} -ben, melyek itt megegyeznek, de nem azonosak. A két rész szimmetriája miatt feltehető, hogy a keresendő két sor az első félben egyezik. Könnyen áttekinthető, hogy ekkor a második félben minden \mathcal{K}_1^{-1} -belihez található két olyan sor, ami ebben egyezik, hozzá tartozó első fele pedig egyforma. Ezzel a konstrukció összes szükséges tulajdonságát igazoltuk.

4. Telített Sperner-rendszerek és lezárások

Egy *Sperner-rendszert* akkor nevezünk *telítettnek*, ha a rendszer nem bővíthető további halmazzal úgy, hogy a *Sperner-tulajdonsága* megmaradjon. Ha \mathcal{L} egy lezárás az X alaphalmazon, a

$$\{K: \mathcal{L}(K) = X, \exists K' \subset K, \mathcal{L}(K') = X\}$$

rendszert, \mathcal{L} minimális kulcsainak rendszerét $\mathcal{K}(\mathcal{L})$ -lél jelöljük.

4.1. TÉTEL. Ha \mathcal{K} egy telített *Sperner-rendszer*, akkor a $\mathcal{K}(\mathcal{L}) = \mathcal{K}$ egyenlőség egyértelműen meghatározza az \mathcal{K} lezárást.

Bizonyítás. Igazoljuk, hogy ha \mathcal{K} telített és $\mathcal{K}(\mathcal{L}) = \mathcal{K}$, akkor minden A esetén meg van határozva. \mathcal{K} telítettsége miatt minden A halmazhoz van egy $K \in \mathcal{K}$, amelyre 1. $A \supseteq K$ vagy 2. $A \subset K$, $A \neq K$ teljesül. Az első esetben $\mathcal{L}(A) = X$, tehát a lezárás valóban egyértelmű. A második esetben azt fogjuk belátni, hogy $\mathcal{L}(A) = A$. Tegyük fel indirekte, hogy $a \in \mathcal{L}(A) - A$. A telítettség miatt ismét van egy $K_1 \in \mathcal{K}$ halmaz, amire vagy 2/a) $A \cup \{a\} \subseteq K_1$ vagy 2/b) $A \cup \{a\} \supset K_1$, $A \cup \{a\} \neq K_1$ teljesül. A 2/a) esetben (2.2) alapján $\mathcal{L}(K_1 - \{a\}) \supseteq K_1 - \{a\}$, (2.3) alapján (mivel $K_1 - \{a\} \supseteq A / \mathcal{L}(K_1 - \{a\}) \supseteq \mathcal{L}(A) \supseteq A \cup \{a\}$ állítható. Összegezve $\mathcal{L}(K_1 - \{a\}) \supseteq K_1$. Amiből (2.3a)-val $\mathcal{L}(K_1 - \{a\}) = \mathcal{L}((K_1 - \{a\})) \supseteq \mathcal{L}(K_1) = X$ következne, ellentmondásban K_1 minimalitásával. A 2/b) esetben $\mathcal{L}(A) \supseteq A \cup \{a\} \supseteq K_1$, vagyis $\mathcal{L}(A) = \mathcal{L}(\mathcal{L}(A)) \supseteq \mathcal{L}(K_1) = X$, ismét ellentmondással. Tehát egy \mathcal{K} telített *Sperner-rendszer* egyértelműen határozza meg \mathcal{L} -et.

A tétel megfordítása azonban nem igaz, mint azt a következő példa mutatja. Legyen az alaphalmaz, X 5-elemű és $\mathcal{K} = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{1, 5\}\}$. Igazoljuk, hogy — bár \mathcal{K} nem telített — egyértelműen meghatározza minden halmaz lezárását. Minden 2-nél több elemű halmaz tartalmaz egy \mathcal{K} -beli halmazt, ezek lezárása tehát X . A \mathcal{K} -beli 2-eleműek lezárása X . A \mathcal{K} -ban nem szereplő 2-eleműek lezárása önmaga. Ugyanis, ha pl. $\mathcal{L}(\{1, 3\})$ legalább 3-elemű volna, akkor tartalmazna egy \mathcal{K} -belit, így $\mathcal{L}(\mathcal{L}(\{1, 3\})) = \mathcal{L}(\{1, 3\})$ tartalmazná X -et, azaz $\mathcal{L}(\{1, 3\}) = X$ lenne. Mivel sem $\{1, 3\}$, sem valamilyen része nincs \mathcal{K} -ban, ellentmondásra jutottunk. Ha most $\mathcal{L}(\{1\})$ -nek volna 1-től különböző eleme, akkor azt $\mathcal{L}(\{1, 3\})$ és $\mathcal{L}(\{1, 4\})$ is tartalmazná. Egyikük ellentmondást okoz. Hasonlóan látható, hogy minden egyelemű halmaz lezárása önmaga, és $\mathcal{L}(\{\emptyset\}) = \emptyset$.

Végtelen sok ilyen példát találtunk, de nem sikerült pontosan jellemeznünk azon \mathcal{K} -kat, amelyek egyértelműen meghatároznak egy lezárást¹.

IRODALOM

- [1] DEMETROVICS, J. and KATONA, G. O. H., "Extremal combinatorical problems in relational data base", *Fundamentals of Computation Th., Lecture Notes in Computer Science* 117 (Springer Verlag, Berlin, Heidelberg, N. Y., 1981) 110—119.
- [2] DEMETROVICS, J. és KATONA, GY., „A relációs adatbázis extrémális problémái”, *Alkalmazott Matematikai Lapok* 8 (1982) 183—193.
- [3] CODD, E. F., "A relational model of data for large shared data banks", *Comm. ACM* 13 (1970) 377—387.
- [4] DEMETROVICS, J., "On the equivalence of candidate keys with Sperner systems", *Acta Cybernetica* 4 (1979) 247—252.
- [5] DEMETROVICS, J. and GYEPESI, GY., "On the functional dependency and some generalization of it", *Acta Cybernetica*.

(Beérkezett: 1982. december 15.)

DEMETROVICS JÁNOS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1132 BUDAPEST, VICTOR HUGO U. 18.

FÜREDI ZOLTÁN ÉS KATONA GYULA

MTA MATEMATIKAI KUTATÓ INTÉZET
1053 BUDAPEST, RÉALTANODA U. 13—15.

¹ Megjegyzés a korrektúránd. Azóta BUROSCH G.-vel közösen sikerült jellemeznünk azon \mathcal{K} -kat, melyek egyértelműen meghatároznak egy lezárást.

THE RELATION BETWEEN THE NUMBER OF INDIVIDUALS
AND DEPENDENCIES IN THE COMPOSED DATA BASES

J. DEMETROVICS, Z. FÜREDI and G. O. H. KATONA

In an earlier paper we have investigated the problem that how many is the minimal number of individuals in a data base with a given key system. In the present paper we investigate the same question and a similar one. Namely we shall determine the minimal number of individuals in that case when in the data base a system of dependencies is also given.

In many cases it is very useful to determine some parameters of a composed data base from the same parameters of the sub data bases. We try to do this for the minimal number of the individuals. We can do this only if the data base is composed from independent parts.

MINIMÁLIS RELÁCIÓS ADATBÁZISOK

FÜREDI ZOLTÁN

Budapest

Jelölje \mathcal{F}_k^n az n -elemű halmaz összes k -elemű részhalmazának rendszerét. Legyen M egy olyan mátrix, amely \mathcal{F}_k^n -t indukálja minimális kulcsrendszerként. DEMETROVICS és KATONA [7] belátták, hogy ekkor az M sorinak minimális száma $s(\mathcal{F}_k^n) > \sqrt{2 \binom{n}{k-1}}$ és, hogy ez a becslés pontos a $k=1$ és 2 esetben. Ismeretes, hogy $s(\mathcal{F}_1^n) = n + O(1)$ (l. [9]).

E dolgozat célja, hogy alkalmas M mátrix megadásával megmutassa, hogy általában is rögzített k mellett

$$s(\mathcal{F}_k^n) = O(n^{(k-1)/2}).$$

1. Alapfogalmak

Az itt következő relációs adatbázis modellt CODD [4], ARMSTRONG [1] és DEMETROVICS [5] dolgozták ki. Egy összefoglaló dolgozatot l. pl. DEMETROVICS—KATONA [7] vagy (magyar nyelven) BÉKÉSSY—DEMETROVICS—GYEPESI [3]. Itt csak a legszükségesebb definíciókra szorítkozunk.

Jelöljön M egy $n \times m$ -es mátrixot (adatbázist). Az oszlopok halmazát X -szel, a sorokét V -vel fogjuk jelölni. ($|X|=n$, $|V|=m$). Egy $\kappa \subset X$ részhalmazt *kulcs*-nak nevezünk, ha a κ oszlopai által alkotott részmátrix minden sora különböző. A κ kulcs *minimális*, ha nem tartalmaz további kulcsot ($\nexists \kappa' \subset \kappa$, $\kappa' \neq \kappa$, κ' kulcs). Az M által indukált minimális kulcsok rendszerét $\mathcal{K}(M)$ -mel, vagy röviden \mathcal{K} -val jelöljük. Nyilván \mathcal{K} -ban nem létezik két egymást tartalmazó halmaz, más néven \mathcal{K} egy *Sperner-rendszer*.

1.1. ÁLLÍTÁS (l. pl. [2]). Tetszőleges \mathcal{F} *Sperner-rendszer*hez létezik olyan mátrix, amely indukálja. (Azaz $\exists M$, melyre $\mathcal{F} = \mathcal{K}(M)$.)

1.2. KÖVETKEZMÉNY. n oszlop esetén a minimális kulcsok maximális száma éppen $\binom{n}{2}$.

2. Minimális indukáló adatbázisok

Legyen \mathcal{F} tetszőleges *Sperter-rendszer* n ponton. Az 1.1. állítás szerint létezik olyan M mátrix, amely indukálja. Természetesen felvetődő kérdés, hogy mennyi ennek minimális sorszáma. Legyen

$$s(\mathcal{F}) = \min \{m: \mathcal{F} = \mathcal{K}(M), \quad M \text{ } n \times m\text{-es mátrix}\}.$$

Vagyis s azt a minimális sorszámot jelöli, ahány sorú mátrixszal \mathcal{F} indukálható.

Jelölje $\Delta\mathcal{F}$ X azon maximális részhalmazait, melyek nem kulcsok. Azaz $\Delta\mathcal{F} = \{G \subset X : G \text{ nem kulcs, de } \forall G' \subsetneq G \subset X \text{ esetén } G' \text{ kulcs}\}$. Az \mathcal{F} és $\Delta\mathcal{F}$ halmazrendszerek egyértelműen meghatározzák egymást.

$$2.1. \text{ ÁLLÍTÁS (l. pl. [2], } \binom{s(\mathcal{F})}{2} \equiv |\Delta\mathcal{F}| \equiv s(\mathcal{F}) - 1.$$

A *Sperner-tétel* és a 2.1. állítás következménye a következő állítás.

2.2. ÁLLÍTÁS. Ha \mathcal{F} egy *Sperner-rendszer* n ponton, akkor

$$s(\mathcal{F}) \leq \left\lfloor \binom{n}{\lfloor \frac{n}{2} \rfloor} \right\rfloor + 1.$$

A következő eredmény, melyet bizonyítás nélkül közlünk, alsó becslést ad a lehetséges maximális $s(\mathcal{F})$ értékre (DEMETROVICS—GYEPESI [6]):

2.3. TÉTEL. Létezik olyan \mathcal{F} *Sperner-rendszer* n ponton, melyre

$$\frac{1}{n^2} \left\lfloor \binom{n}{\lfloor \frac{n}{2} \rfloor} \right\rfloor \leq s(\mathcal{F}).$$

3. Uniform rendszereket indukáló adatbázisok

A 2.3. tétel bizonyítása (l. még [3] és [7]) nem konstruktív. A konkrét konstrukciók keresése vezethet az uniform rendszerek vizsgálatához. Jelölje \mathcal{F}_k^n az n -elemű alaphalmaz összes k -elemű részhalmazából álló halmazrendszert. Nyilván $\Delta\mathcal{F}_k^n = \mathcal{F}_{k-1}^n$, így a 2.1. állításból kapjuk, hogy

$$3.1. \text{ KÖVETKEZMÉNY. } \sqrt{2 \binom{n}{k-1}} < s(\mathcal{F}_k^n) \leq \binom{n}{k-1} + 1.$$

DEMETROVICS és KATONA [7] meghatározták $s(\mathcal{F}_k^n)$ értékét, ha $k=1, 2, n-1, n$.

$$3.2. \text{ TÉTEL. } s(\mathcal{F}_1^n) = 2, \quad s(\mathcal{F}_2^n) = \lfloor (1 + \sqrt{8n+1})/2 \rfloor \quad \text{és}$$

$$s(\mathcal{F}_n^n) = n, \quad s(\mathcal{F}_n^n) = n + 1.$$

Továbbá bebizonyították, hogy

$$(3.1) \quad n \leq s(\mathcal{F}_3^n) \leq \mathcal{O}(n^{3/2}).$$

A felső becslés a nyilvánvaló

$$(3.2) \quad s(\mathcal{F}_k^n) \leq s(\mathcal{F}_k^{n-1}) + s(\mathcal{F}_{k-1}^{n-1})$$

összefüggésből adódik.

Legutóbb feloldható blokkrendszerek segítségével sikerült bizonyítani sejtésüket, miszerint $s(\mathcal{F}_3^n) = n + \mathcal{O}(1)$.

$$3.3. \text{ TÉTEL (l. [8]). } n \leq s(\mathcal{F}_3^n) \leq n + 8.$$

4. Eredmények

A fent említett eredmények közül azokra koncentrálunk, melyekben a kulcsok nagysága rögzített. A 3.2. és a 3.3. tétel azt mutatja, hogy a 3.1. következmény alsó becslése adja $s(\mathcal{F}_k^n)$ pontos nagyságrendjét (legalábbis a $k=1, 2$ és 3 esetekben). E dolgozat célja, hogy alkalmas M mátrix megadásával bizonyítsa, hogy ez minden k -ra így van, azaz

4.1. TÉTEL. Ha $n > k \geq 1$ egészek, akkor $s(\mathcal{F}_k^n) < n^{(k-1)/2} \cdot 2^k$.

Felmerülhet 2.3 tétellel analóg kérdés a k -uniform kulcsrendszerek esetén is.

4.2. TÉTEL. (1. [8]). Létezik olyan $\mathcal{H} \subset \mathcal{F}_k^n$ melyre $s(\mathcal{H}) > \sqrt{2} \binom{2k-2}{k-1}^{1/n/(2k-2)/2}$.

Jelölje $f_k(n) = \max \{s(\mathcal{F}): \mathcal{F} \subset \mathcal{F}_k^n\}$. Könnyű ellenőrizni, hogy $f_1(n) = 2$. A fenti tétel adja, hogy $f_k(n) > 2^{n/4}$. Kissé talán meglepő, hogy egy egyszerű konstrukció is csak ilyen sok sorral imprikálható. Ezzel kapcsolatos a következő probléma:

4.3. PROBLÉMA. Határozzuk meg $\max \{|\Delta \mathcal{F}|: \mathcal{F} \subset \mathcal{F}_k^n\}$ -et. Más szóval: Egy n pontú k -gráfnak hány maximális független pontrendszere lehet? A $k=2$ eset megoldotj: diszjunkt háromszögek rendszere $3^{\lfloor n/3 \rfloor}$ -ot ad.

A $\max s(\mathcal{H})$ meghatározása, a minimális sorszámra alsó becslést nyújtó technikák híján, nagyon nehéznek látszik.

5. A konstrukció előkészítése, véges testek és differenciahalmazok

A 4.1. tétel bizonyításához szükség lesz néhány eddig nem szerepelt fogalomra.

Jelöljön Q egy véges q -elemű testet. (Pl. ha q prím, akkor $Q = \langle 0, 1, \dots, q-1 \rangle$ a modulo q -vett szorzással és összeadással.) Az $A \subset Q$ részhalmazt *differenciahalmaznak* mondjuk, ha két A -beli elem különbségeként minden testelem megkapható. Azaz, ha az $x = a_i - a_j$ egyenlet megoldható minden $x \in Q$ esetén ($a_i, a_j \in A$).

5.1. LEMMA. A q -elemű véges test tartalmaz egy legfeljebb $2\sqrt{q}$ tagú differenciahalmazt.

Bizonyítás. Legyen e a Q egységeleme és $a = \lceil \sqrt{q} \rceil$. Végül legyen $A = \{0, e, 2e, \dots, ae\} \cup \{2ae, 3ae, \dots, a^2e\}$. Ez az A nyilván differenciahalmaz, hiszen minden $(q-a)$ -nál kisebb x egész felírható $x = \alpha a - \beta$ alakban, ahol $0 \leq \beta < a$ és $\alpha a < q$. Végül a $q-a$ és $q-1$ között egészek is előállíthatóak (modulo q) $0 - \beta$ alakban. Ennek az A -nak a számosságára:

$$|A| = a + \left\lfloor \frac{q}{a} \right\rfloor = \lceil \sqrt{q} \rceil + \left\lfloor \frac{q}{\lceil \sqrt{q} \rceil} \right\rfloor \leq \lceil \sqrt{q} \rceil + \lfloor \sqrt{q} \rfloor \leq \lceil 2\sqrt{q} \rceil.$$

(Ugyanis minden $x \geq 0$ valósra $\lceil x \rceil + \lfloor x \rfloor \leq \lceil 2x \rceil$.)

Végezetül könnyen látható, hogy ha \sqrt{q} nem egész szám, akkor az $A - \{0\}$ is differenciahalmaz. Ezzel a lemmát igazoltuk.

5.2. *Megjegyzés.* Véges test helyett a differenciahalmaz véges csoportra is értelmezhető. Mi itt tulajdonképpen azt bizonyítottuk, hogy a q -elemű véges ciklikus csoport minimális differenciahalmaza legfeljebb $2\sqrt{q}$ elemű. Általános tételt (amely minden csoportra igaz) nem ismerek, de a *mohó algoritmussal* (vagy LOVÁSZ LÁSZLÓ [11] egy tételének következményeként) könnyen igazolható, hogy

(5.1) tetszőleges q -elemű véges csoport minimális differenciahalmaza legfeljebb $q/\log q$ elemű.

Azt hiszem az (5.1) általában nem javítható.

Az itt ismertetendő konstrukciónk csupán véges gyűrűket használ, így a fenti megjegyzésnek egy pontosabb felső becslés kidolgozásában lehet jelentősége.

5.3. *Megjegyzés.* Az 5.1. lemma egyszerűsége ellenére azért is jól használható, mert bizonyos értelemben elég közel van az igazsághoz. (Konstans szorzó erejéig.) Ugyanis

(5.2) Tetszőleges $A \subset Q$ differenciahalmazra $|A| \geq \sqrt{q}$.

Valóban az összes $a_i - a_j$ párok száma éppen $|A|^2$ ($a_i, a_j \in A$), így ha ezek kiadják a Q összes elemét, akkor $|A|^2 \geq |Q|$. (Ráadásul e párok nem is mind különbözőek.)

5.4. *Megjegyzés.* A számelméletben és egyes algebrai struktúrákban sokat használt a *perfekt* differenciahalmaz fogalma. Ez olyan differenciahalmaz, mely az $x \neq 0$ -ra az $x \in Q$ elemet egyértelműen állítja elő $a_i - a_j$ alakban. Ilyen perfekt differenciahalmaz csak a $q = n^2 + n + 1$ esetén létezik (ahol n egész szám).

Bővebbet erről l. [10].

6. A 4.1. tétel bizonyítása, a konstrukció

Legyen $A \subset Q$ a q -elemű véges test egy minimális számosságú differenciahalmaza, mely tartalmazza a 0-t és az 1-et. Az 5.1. lemma szerint $|A| = a < 2\sqrt{q}$. Definálunk egy $q \times 2a^{k-1}$ -es M mátrixot, melynek elemei testelemek lesznek és egy-egy sora egy-egy legfeljebb $(k-1)$ -edfokú Q feletti polinom, $0, 1, \dots, g-1$ helyen felvett helyettesítési értékeiből áll. Legyen a \mathcal{P} polinomhalmaz a következő:

$$\mathcal{P} = \{c_{k-1}x^{k-1} + c_{k-2}x^{k-2} + \dots + c_1x + c_0 : c_{k-1} = 0 \text{ vagy } 1, \text{ és minden } c_i \in A\}.$$

Legyen M egy $Q \times \mathcal{P}$ -es mátrix, ahol a p -edik sor x -edik eleme $p(x)$ ($p \in \mathcal{P}$, $x, p(x) \in Q$). Belátjuk, hogy ez az M rendelkezik a kívánt tulajdonságokkal.

6.1. **LEMMA.** Az M mátrix p' és p'' -edik sorának közös elemeinek halmaza éppen a $(p' - p'')$ polinom zérushelyeinek halmaza.

A 6.1. lemma bizonyítása triviális, hiszen ahol $p'(x) = p''(x)$, ott $(p' - p'')(x) = 0$.

6.2. **KÖVETKEZMÉNY.** Az összes legalább k -elemű halmaz kulcs.

A 6.2. következmény bizonyítása. Itt használjuk, hogy a Q véges test és ezért a Q feletti legfeljebb $(k-1)$ -edfokú polinomoknak legfeljebb $(k-1)$ gyökük lehet.

6.3. LEMMA. Ha a Q feletti p polinom legfeljebb $(k-1)$ -edfokú és fő együtt-hatója 1, akkor felírható $p'-p''$ alakban, ahol $p', p'' \in \mathcal{P}$.

A 6.3. lemma bizonyítása a differenciahalmaz fogalmából adódik, hiszen ha $p(x) = c_{k-1}x^{k-1} + \dots + c_1x + c_0$, akkor az összes c_i -t szézbontva $a'_i - a''_i$ alakra ($a'_i, a''_i \in A$) a $p'(x) = \sum_{i \leq k-1} a'_i x^i$, $p''(x) = \sum_{i \leq k-1} a''_i x^i$ polinomok kielégítik a lemma feltételeit.

6.4. KÖVETKEZMÉNY. Az M által indukált kulcsok között nincs egyetlen k -nál kisebb elemű.

A 6.4. következmény bizonyítása. Legyen $T \subset Q$ $|T| \leq k-1$, $T = \langle t_1, \dots, t_i, \dots \rangle$. Tekintsük a $p_T = (x-t_1)(x-t_2)\dots(x-t_i)\dots$ legfeljebb $(k-1)$ -edfokú 1 főegyütt-hatójú polinomot. Ehhez található a 6.3. lemma szerint egy olyan p' és $p'' \in \mathcal{P}$, melyre $p'-p'' = p_T$. De ekkor a 6.1. lemma szerint a p' és p'' -edik sorok a T halmazban megegyeznek, így T nem lehet M által indukált kulcs.

Összefoglalva

6.5. TÉTEL. Ha $n > k \geq 1$, akkor a fent definiált M -re $\mathcal{K}(M) = \mathcal{F}_k^Q$.

A 4.1. tétel bizonyítása. Nyilvánvaló, hogy ha Q oszlopaiból elhagyunk néhányat, akkor (a megmaradt oszlopok halmazát Y -nal jelölve)

$$(6.1) \quad \mathcal{K}(M|Y) = \mathcal{F}_k^{Q|Y}.$$

Ebből a Csebisev-tételt használva (mely szerint minden n és $2n$ között létezik prím-szám) legyen $n \leq q \leq 2n$ adódik, hogy

$$s(\mathcal{F}_k^n) \leq s(\mathcal{F}_k^{Q/n}) \leq s(\mathcal{F}_k^Q) = 2a^{k-1} \leq 2(2q)^{\frac{k-1}{2}} < 2(4n)^{\frac{k-1}{2}} = 2^k n^{(k-1)/2}.$$

7. További bizonyítások

A 4.2. tétel bizonyítása szintén alkalmas konstrukció megadásával történik. Particionáljuk X -et, $X = X_1 \cup \dots \cup X_q \cup Y$, ahol $q = \lfloor n/(2k-2) \rfloor$ és $|X_i| = 2k-2$ ($1 \leq i \leq q$). Legyen $\mathcal{F} = \{F: |F| = k, F \subset X_i \text{ valamely } i\text{-re}\}$. Ekkor $\Delta\mathcal{F} = \{A: |A \cap X_i| = k-1 \text{ minden } i\text{-re, és } Y \subset A\}$. Következésképp $|\Delta\mathcal{F}| = \binom{2k-2}{k-2}^q$. Végül a 2.1 állítást használva kap-hatun alsó becslést $s(\mathcal{F})$ -re.

IRODALOM

- [1] ARMSTRONG, W. W., "Dependency structures of data base relationship", *Information Processing* 74 (North-Holland Publishing Company, Amsterdam, 1974) 580—583.
- [2] BÉKÉSSY, A., DEMETROVICS, J., HANNÁK, L., KATONA, G. O. H. and FRANKL, P., "On the number of maximal dependencies in a data relation of fixed order" *Discrete Mathematics* 30 (1980) 83—88.
- [3] BÉKÉSSY, A., DEMETROVICS, J. és GYEPESI, GY., „Relációs adatbázis logikai szintű vizsgálata funkcionális függőségek szempontjából”, *MTA SZTAKI Tanulmányok* 109 (1980) 1—56.

- [4] CODD, E. F., "A relational model of data for large shared data banks", *Communications of the ACM* 13 (1970) 377—387.
- [5] DEMETROVICS, J., "Candidate keys and antichains", *SIAM Journal on Algebra and Discrete Methods* 1 (1980) 1—92.
- [6] DEMETROVICS, J. and GYEPESI, GY., "On the functional dependency and some generalization of it", *Acta Cybernetica* 5 (1981) 296—305.
- [7] DEMETROVICS, J. and KATONA, G. O. H., "Extremal combinatorial problems in relational data base", *Lecture Notes in Computer Sciences* 117 (1981) 110—119.
- [8] DEMETROVICS, J., FÜREDI, Z. and KATONA, G. O. H., "Maximum matrix representation of closure operations", *Discrete Mathematics*, megjelenés alatt.
- [9] FÜREDI, Z., „Relációs adatbázisokkal kapcsolatos kombinatorikai problémák”, kézirat, Budapest, 1981.
- [10] HALBERSTAM, H. and ROTH, F. K., *Sequences* (Clarendon Press, Oxford, 1966).
- [11] LOVÁSZ, L., Minimax tételek a kombinatorikában, Doktori értekezés, Budapest, 1978.
- [12] SPERNER, E., „Ein Satz über Untermengen einer endlich Menge“, *Mathematische Zeitschrift* 27 (1928) 544—548.

(Beérkezett: 1982. december 15.)

FÜREDI ZOLTÁN

MTA MATEMATIKAI KUTATÓ INTÉZET
1053 BUDAPEST, RÉALTANODA U. 13—15.

MINIMAL RELATIONAL DATA BASES

Z. FÜREDI

Let F_k^n denote the system of all k -element subsets of an n -element set. Let \mathbf{M} be a matrix which generates F_k^n as a minimal key system. DEMETROVICS and KATONA have proved ([7]) that in this case the minimal number of the rows of \mathbf{M} is $s(F_k^n) > \sqrt{2 \binom{n}{k-1}}$ and this bound is exact for $k=1$ and 2. It is also well known that $s(F_2^n) = n + O(1)$ (see [8]).

The purpose of this paper is to show that generally for a fixed value of k

$$s(F_k^n) = O(n^{(k-1)/2}).$$

This is done by definition of an appropriate matrix \mathbf{M} .

EGYLÉPÉSES MÓDSZERCSALÁDOK KONVERGENCIA- ÉS HIBAANALÍZISE

GALÁNTAI AURÉL

Gödöllő

A dolgozatban egylépéses módszerekből álló összetett módszerek konvergenciáját és hibaterjedési viszonyait vizsgáljuk.

1. Bevezetés

Tekintsük az

$$(1.1) \quad y' = f(t, y); \quad y(t_0) = y_0 \quad (f \in C([t_0, b] \times R^m, R^m))$$

alakú *Cauchy-problémákat*. Ismeretes, hogy nincs minden ilyen feladatra egyformán jó numerikus módszer, sőt a gyakorlatban felmerülő különböző problémátípusok hatékony megoldási módszerei egymástól erősen eltérő tulajdonságúak lehetnek. Ez a probléma egy feladat numerikus megoldásának a során is felléphet, ha egy elég hosszú számítási intervallumon belül az elméleti megoldás jellege (pl. a megoldás bizonyos deriváltjainak normája) erősen megváltozik.

Abból a célból, hogy minél szélesebb feladatosztályokon kielégítő hatékony-ságú (ún. robusztus) numerikus eljárásokat nyerjünk, az irodalomban bevezették a több módszerből álló komplex módszercsaládokat. Ilyenek például a ciklikus előre lépő módszerek (lásd STETTER [16]), vagy a változólépéshosszú-változó formulájú módszerek (ZLATEV [19]). A leghíresebb komplex módszert és ennek FORTRAN programját GEAR fejlesztette ki 1971-ben ([6]). E módszer igen sikeresnek bizonyult az (1.1) alakú feladatok hatékony megoldásában, ezért számos változata, például GEAR, EPISODE, DESOL ([12]) ismeretes. GEAR eljárása retrográd differencia formulákból és *Adams-típusú módszerekből*, valamint a lépéshossz, a formula és a rend választását szabályozó döntési kritériumokból áll. Az ilyen típusú eljárások konvergenciáját GEAR, TU és WATANABE 1974-ben bizonyították be ([7], [8]).

Az irodalomban ismeretesek egylépéses formulákból összeállított módszer-családok is ([9], [15]), azonban az ezirányú vizsgálatok főleg experimentálisak. A dolgozatban igazoljuk az egylépéses módszercsaládok konvergenciáját és becsüljük a konvergencia sebességét. Igazoljuk továbbá a lokális hibabecslési elv helyességét is.

2. Alapfogalmak, jelölések

Tegyük fel, hogy létezik $L \geq 0$ konstans úgy, hogy

$$(2.1) \quad \|f(t, y) - f(t, y^*)\| \leq L \|y - y^*\| \quad (t \in [t_0, b]; \quad y, y^* \in R^m).$$

Tegyük fel továbbá, hogy minden $t^* \in [t_0, b]$ és minden $y^* \in R^m$ esetén az

$$(2.2) \quad y' = f(t, y); \quad y(t^*) = y^*$$

Cauchy-feladatnak pontosan egy, a $[t_0, b]$ intervallumon értelmezett megoldása van. Legyen $t_0 < x \leq b$ tetszőleges rögzített pont és jelöljön Δ_N tetszőleges, a

$$\Delta_N: t_0 < t_1 < \dots < t_N = x$$

feltételnek eleget tevő felosztást. A $[t_0, x]$ intervallum összes felosztásainak halmazát jelöljük π_x -szel. A $\Delta_N \in \pi_x$ felosztás normáját $\|\Delta_N\| = \max \{t_{i+1} - t_i \mid i = 0, 1, \dots, N-1\}$ definiálja, ahol $h_i = t_{i+1} - t_i$ az i -edik lépéshossz. A *Cauchy-probléma* $t_n \in \Delta_N$ pontbeli közelítő megoldását jelölje y_n .

Legyen $W = \{1, 2, \dots, w^*\}$ véges és tegyük fel, hogy minden $w \in W$ index esetén az

$$(2.3) \quad y_{n+1} - y_n = h_n \psi(w|t_n, y_n, y_{n+1}, h_n) \quad (t_{n+1} \in \Delta_N)$$

egylépéses módszer *Henrici-féle értelemben konvergens*. Ez azt jelenti (lásd [4]), hogy minden $w \in W$ index esetén a

$$\psi(w|., \dots, .) \in C([t_0, b] \times R^m \times R^m \times [0, b - t_0], R^m)$$

növekmény függvény kielégíti a

$$(2.4) \quad \psi(w|t, y, y, 0) \equiv f(t, y) \quad (t \in [t_0, b]; y \in R^m)$$

feltételt, továbbá létezik olyan $K_w \geq 0$ konstans, amellyel fennáll

$$(2.5) \quad \|\psi(w|t, y, z, h) - \psi(w|t, y^*, z^*, h)\| \leq K_w (\|y - y^*\| + \|z - z^*\|)$$

minden $t \in [t_0, b]$, $h \in [0, b - t_0]$ és $y, y^*, z, z^* \in R^m$ esetén.

2.1. DEFINÍCIÓ. Legyen $I = I(\Delta_N): \{0, 1, \dots, N-1\} \rightarrow W$ tetszőleges indexfüggvény. A (ψ, W, I) hármast egylépéses módszer családnak nevezzük, ha az y_n ($n = 0, 1, \dots, N$) közelítő megoldást az

$$(2.6) \quad y_{n+1} - y_n = h_n \psi(I(n)|t_n, y_n, y_{n+1}, h_n) \quad (t_{n+1} \in \Delta_N)$$

rekurzió definiálja.

Ez a definíció lényegében azt jelenti, hogy a számítások minden egyes lépésében az I indexfüggvény határozza meg az aktuálisan használt formulát. A h_n lépéshossz az $I(n)$ értéktől függetlenül tetszőlegesen változtatható.

2.2. DEFINÍCIÓ. A (ψ, W, I) módszer családot konvergensnek nevezzük, ha minden $x \in [t_0, b]$ esetén teljesül, hogy

$$(2.7) \quad \max_{t_n \in \Delta_N} \|y_n - y(t_n)\| \rightarrow 0 \quad (\Delta_N \in \pi_x, N \rightarrow +\infty, \|\Delta_N\| \rightarrow 0).$$

Szükséges még a következő

2.3. DEFINÍCIÓ. Jelölje $Y_j: [t_0, b] \rightarrow R^m$ a perturbált

$$(2.8) \quad y' = f(t, y); \quad y(t_j) = y_j \quad (j = 0, 1, \dots, N)$$

Cauchy-probléma elméleti megoldását. A (ψ, W, I) módszercsalád (2.8) problémára vonatkozó $t_n \in \Delta_N$ pontbeli lokális hibáját a

$$(2.9) \quad T_j(I(n)|t_n, h_n) = Y_j(t_n) + h_n \psi(I(n)|t_n, Y_j(t_n), Y_j(t_{n+1}), h_n) - Y_j(t_{n+1})$$

képlet definiálja.

A $j=0$ esetben az egyszerűsítő $T(I(n)|t_n, h_n)$ jelölést használjuk, mert $Y_0(t) \equiv y(t)$.

3. A konvergencia vizsgálata

Először a (ψ, W, I) módszercsalád stabilitását igazoljuk. Ehhez tekintsük az

$$(3.1) \quad y_{n+1}^* - y_n^* = h_n \psi(I(n)|t_n, y_n^*, y_{n+1}^*, h_n) + d_{n+1} (t_{n+1} \in \Delta_N)$$

perturbált rekurziót tetszőleges $\{d_n\}_{n=1}^N \subset R^m$ perturbáció esetén. Legyen továbbá $r_n = y_n^* - y_n$ ($n=0, 1, \dots, N$). Ekkor igaz a

3.1. TÉTEL. Tetszőleges (ψ, W, I) módszercsaládhoz létezik olyan $c_1, c_2 > 0$ konstans, amelyekkel $\Delta_N \in \pi_x$ és $\|\Delta_N\| \leq 1/2K^*$ ($K^* = \max\{K_w | w \in W\}$) esetén az y_n, y_n^* ($n=0, 1, \dots, N$) közelítő megoldások kielégítik a

$$(3.2) \quad c_1 \left(\|r_0\| + \max_{1 \leq k \leq N} \left\| \sum_{i=1}^k d_i \right\| \right) \leq \max_{0 \leq n \leq N} \|y_n - y_n^*\| \leq \\ \leq c_2 \left(\|r_0\| + \max_{1 \leq k \leq N} \left\| \sum_{i=1}^k d_i \right\| \right)$$

kétoldali egyenlőtlenséget. A konstansok értékéül a

$$c_1 = [3 + 2K^*(b - t_0)]^{-1}, \quad c_2 = 2[1 + 4K^*(b - t_0)e^{4K^*(b - t_0)}]$$

számok választhatók.

Az állítás bizonyításához szükségünk van a *Gronwall—Bellman-lemma* diszkrét változatára (lásd pl. [13]).

LEMMA. Tegyük fel, hogy az $x(i)$, $y(i)$ és $z(i)$ számok ($i=0, 1, \dots, n$) kielégítik a $z(i) \geq 0$ és az

$$(3.3) \quad x(i) \leq y(i) + \sum_{j=0}^{i-1} z(j)x(j) \quad (i = 0, 1, \dots, n)$$

feltételeket. Ekkor igaz, hogy

$$(3.4) \quad x(n) \leq y(n) + \sum_{i=0}^{n-1} z(i)y(i) \cdot \prod_{j=i+1}^{n-1} [1 + z(j)]$$

A 3.1. tétel bizonyítása. A (2.6) és (3.1) egyenlet kivonásából kapjuk, hogy

$$(3.5) \quad r_{n+1} = r_n + h_n \{ \psi(I(n)|t_n, y_n^*, y_{n+1}^*, h_n) - \psi(I(n)|t_n, y_n, y_{n+1}, h_n) \} + d_{n+1},$$

ahonnan összegezéssel adódik, hogy

$$(3.6) \quad r_{n+1} = r_0 + \sum_{i=0}^n h_i \{ \psi(I(i)|t_i, y_i^*, y_{i+1}^*, h_i) - \psi(I(i)|t_i, y_i, y_{i+1}, h_i) \} + \sum_{i=0}^n d_{i+1}.$$

Innen a (2.5) egyenlőtlenség alapján adódik, hogy

$$\|\mathbf{r}_{n+1}\| \leq \|\mathbf{r}_0\| + \sum_{i=0}^n h_i K_{I(i)} (\|\mathbf{r}_i\| + \|\mathbf{r}_{i+1}\|) + \left\| \sum_{i=1}^{n+1} \mathbf{d}_i \right\|.$$

A $\|\Delta_N\| K^* \leq 1/2$ és az $(1-x)^{-1} \leq 2$ ($0 \leq x \leq 1/2$) összefüggések felhasználásával kapjuk, hogy

$$\|\mathbf{r}_{n+1}\| \leq 2\|\mathbf{r}_0\| + 2 \left\| \sum_{i=1}^{n+1} \mathbf{d}_i \right\| + 2K^* \sum_{i=0}^n (h_{i-1} + h_i) \|\mathbf{r}_i\| \quad (h_j := 0, j < 0).$$

A Gronwall—Bellman-lemma és az $1+x \leq e^x$ ($x \geq 0$) egyenlőtlenség alapján

$$\begin{aligned} \|\mathbf{r}_{n+1}\| &\leq 2 \left(\|\mathbf{r}_0\| + \left\| \sum_{i=1}^{n+1} \mathbf{d}_i \right\| \right) + 4K^* \sum_{i=0}^n (h_{i-1} + h_i) \left(\left\| \sum_{j=1}^i \mathbf{d}_j \right\| + \|\mathbf{r}_0\| \right) \\ &\quad \cdot \prod_{s=i+1}^{n-1} [1 + 2K^* (h_{s-1} + h_s)], \end{aligned}$$

illetve

$$\|\mathbf{r}_{n+1}\| \leq 2[1 + 4K^*(b-t_0)e^{4K^*(b-t_0)}] \left\{ \|\mathbf{r}_0\| + \max_{1 \leq k \leq n+1} \left\| \sum_{i=1}^k \mathbf{d}_i \right\| \right\},$$

ami már megfelel az állításban szereplő jobb oldali egyenlőtlenségnek. A bal oldali egyenlőtlenség igazolásához rendezzük át a (3.6) összefüggést a következőképpen

$$\sum_{i=1}^{n+1} \mathbf{d}_i = \mathbf{r}_{n+1} - \mathbf{r}_0 - \sum_{i=0}^n h_i \{ \psi(I(i)|t_i, \mathbf{y}_i^*, \mathbf{y}_{i+1}^*, h_i) - \psi(I(i)|t_i, \mathbf{y}_i, \mathbf{y}_{i+1}, h_i) \},$$

ahonnan egyszerű becslésekkel adódik, hogy

$$\left\| \sum_{i=1}^{n+1} \mathbf{d}_i \right\| \leq \|\mathbf{r}_0\| + \|\mathbf{r}_{n+1}\| + \sum_{i=0}^n h_i K_{I(i)} (\|\mathbf{r}_i\| + \|\mathbf{r}_{i+1}\|) \leq [2 + 2K^*(b-t_0)] \max_{i \leq n+1} \|\mathbf{r}_i\|.$$

Innen a (3.2) becslés bal oldala egyszerű átrendezéssel adódik, amivel a tétel állítását teljes mértékben beláttuk.

Ha $\mathbf{y}_0 = \mathbf{y}_0^*$, $\mathbf{d}_{i+1} = -T(I(i)|t_i, h_i)$ ($i=0, 1, \dots, N-1$) és $\|\Delta_N\| K^* \leq 1/2$ teljesül a $\Delta_N \in \pi_x$ felosztásra, akkor könnyű belátni, hogy

$$(3.7) \quad \mathbf{y}_i^* = \mathbf{y}(t_i) \quad (i = 1, \dots, N).$$

Ily módon a 3.1. tétel speciális eseteként adódik az

1. KÖVETKEZMÉNY. Ha a $\Delta_N \in \pi_x$ felosztás kielégíti a $\|\Delta_N\| \leq 1/2K^*$ feltételt, akkor a (ψ, W, I) módszer család globális hibájára teljesül, hogy

$$\begin{aligned} (3.8) \quad c_1 \max_{0 \leq k \leq N-1} \left\| \sum_{i=0}^k T(I(i)|t_i, h_i) \right\| &\leq \max_{0 \leq n \leq N} \|\mathbf{y}_n - \mathbf{y}(t_n)\| \leq \\ &\leq c_2 \max_{0 \leq k \leq N-1} \left\| \sum_{i=0}^k T(I(i)|t_i, h_i) \right\|. \end{aligned}$$

A következőkben a $T(w|t_i, h_i)h_i^{-1}$ mennyiséget becsüljük rögzített $w \in W$ esetén. A definíció alapján

$$\begin{aligned} \|T(w|t_i, h_i)h_i^{-1}\| &\leq \left\| \frac{y(t_{i+1}) - y(t_i)}{h_i} - \psi(w|t_i, y(t_i), y(t_i), 0) \right\| + \\ &+ \|\psi(w|t_i, y(t_i), y(t_i), 0) - \psi(w|t_i, y(t_i), y(t_i), h_i)\| + \\ &+ \|\psi(w|t_i, y(t_i), y(t_i), h_i) - \psi(w|t_i, y(t_i), y(t_{i+1}), h_i)\|. \end{aligned}$$

A (2.4) összefüggés felhasználásával a jobb oldal első tagjának becslése

$$h_i^{-1} \left\| \int_{t_i}^{t_{i+1}} [f(t, y(t)) - f(t_i, y(t_i))] dt \right\| \leq h_i^{-1} \int_{t_i}^{t_{i+1}} \omega(y'(t); h_i) dt = \omega(y'(t); h_i),$$

ahol $\omega(z(t); \delta)$ a $z \in C([t_0, b], R^m)$ függvény folytonossági modulusát jelöli, azaz az

$$(3.9) \quad \omega(z(t); \delta) = \sup \{ \|z(t) - z(t^*)\| \mid t, t^* \in [t_0, b], |t - t^*| < \delta \}$$

mennyiséget. A jobb oldal második tagját szintén az $\omega(z_w(s); h_i)$ mennyiséggel becsülhetjük felülről, ahol

$$z_w(s) = \max_{t_0 \leq t \leq b} \|\psi(w|t, y(t), y(t), s) - \psi(w|t, y(t), y(t), 0)\|$$

és $z_w \in C([0, b - t_0], R)$. Az utolsó tagot (2.5) alapján

$$K_w \|y(t_i) - y(t_{i+1})\| \leq K_w \omega(y(t); h_i)$$

becsli. E becslések összegzésével kapjuk, hogy

$$(3.10) \quad \|T(w|t_i, h_i)\| \leq h_i [\omega(y'(t); h_i) + \omega(z_w(s); h_i) + K_w \omega(y(t); h_i)].$$

A (3.8) egyenlőtlenség felhasználásával kapjuk, hogy

$$\begin{aligned} \max_{0 \leq n \leq N} \|y_n - y(t_n)\| &\leq c_2 \sum_{w \in W} \sum_{i=0}^{N-1} \|T(w|t_i, h_i)\| \leq \\ &\leq (b - t_0) c_2 \sum_{w \in W} [\omega(y'(t); \|A_N\|) + \omega(z_w(s); \|A_N\|) + K^* \omega(y(t); \|A_N\|)]. \end{aligned}$$

Mint hogy W véges halmaz, a folytonossági modulus ismert tulajdonsága miatt a fenti egyenlőtlenségből következik, hogy

$$(3.11) \quad \max_{0 \leq n \leq N} \|y_n - y(t_n)\| \rightarrow 0 \quad (A_N \in \pi_x, \|A_N\| \rightarrow 0).$$

A fentiek alapján tehát igaz a

3.2. TÉTEL. A (ψ, W, I) egy lépéses módszercsalád minden W indexhalmaz és minden I index függvény esetén konvergens.

A konvergencia sebességét a (3.8) kétoldali becslés pontosan megadja. A könnyebb kezelhetőség érdekében tegyük fel, hogy minden $w \in W$ esetén létezik $p_w, D_w > 0$ konstans, hogy

$$(3.12) \quad \|T(w|t_i, h_i)\| \leq D_w h_i^{p_w+1} \quad (t_{i+1} \in A_N)$$

teljesül. Ekkor (3.8) alapján igaz a

$$(3.13) \quad \max_{0 \leq n \leq N} \|y_n - y(t_n)\| \leq c_2(b-t_0) \left(\max_{w \in W} D_w \right) \|\Delta_N\|^{(\min_{t \in I} p_{t(n)})}$$

egyenlőtlenség.

Az egyenlőtlenség alapján nyilvánvaló, hogy a módszercsalád konvergenciájának a sebességét a $\Delta_N \in \pi_x$ felosztáson előforduló legkisebb p_w rend határozza meg. Ezért a rend csökkentése csak akkor előnyös, ha a megfelelő lokális hibakomponens lényegesen csökken.

Végül megemlítjük a következő speciális esetet. Ha $W = \{1, 2\}$, $p_2 = p_1 + 1$ és $\{\Delta_N\}_{N=1}^\infty \subset \pi_x$, $\|\Delta_N\| \rightarrow 0$ esetén a $k_N = |\{i | I(i) = 1, 0 \leq i < N\}|$ sorozat korlátos, akkor létezik olyan $c_3 > 0$ konstans, hogy

$$\max_{0 \leq n \leq N} \|y_n - y(t_n)\| \leq c_1 \|\Delta_N\|^{p_2}.$$

4. A lokális hibabecslési elv

Az (1.1) problémát általában akkor tekintjük numerikus értelemben megoldottnak, ha az $\{y_n\}_{n=0}^N$ közelítő megoldás kielégíti az

$$(4.1) \quad \|y_n - y(t_n)\| \leq \varepsilon^* \quad (t_n \in \Delta_N)$$

feltételt, ahol $\varepsilon^* > 0$ az előírt hibakorlát.

A (4.1) feltétel ellenőrzését úgy szokás elvégezni, hogy becsüljük a

$$(4.2) \quad T_n(I(n)|t_n, h_n) \quad (n = 0, 1, \dots, N-1)$$

lokális hibákat és ellenőrizzük a

$$(4.3) \quad \|T_n(I(n)|t_n, h_n)\| \leq ch_n \varepsilon^*, \quad [T_n(I(n)|t_n, h_n) \leq c\varepsilon^*]$$

feltételek teljesülését. A $c > 0$ [$c(\Delta_N) > 0$] konstansokat, amelyek az (1.1) problémától és a (ψ, W, I) módszercsaládtól függenek, a gyakorlatban tapasztalati úton választják meg.

Azt a feltételezést, hogy a (4.3) feltételek fennállásából a (4.1) feltétel fennállása következik, lokális hibabecslési elvnek nevezzük (lásd pl. [4], [5]).

A lokális hibabecslési elv megalapozását tartalmazza a

4.1. TÉTEL. Létezik $h^* > 0$ konstans, hogy ha (ψ, W, I) módszercsalád lokális hibájára a $\Delta_N \in \pi_x$ ($\|\Delta_N\| \leq h^*$) felosztáson teljesülnek a

$$(4.4) \quad \|T_n(I(n)|t_n, h_n)\| \leq h_n \varepsilon, \quad [\|T_n(I(n)|t_n, h_n)\| \leq \varepsilon], \quad (n = 0, 1, \dots, N-1)$$

feltételek, akkor

$$(4.5) \quad \max_{0 \leq n \leq N} \|y_n - y(t_n)\| \leq c_1 \varepsilon$$

is fennáll alkalmas $c_3 = c_3(f) > 0$ [$c_3 = c_3(f, \Delta_N) > 0$] konstanssal.

Bizonyítás. A *Peano-egyenlőtlenség* speciális eseteként ismeretes, hogy az $y' = f(t, y)$ differenciálegyenlet bármely két φ_1, φ_2 megoldására teljesül, hogy

$$(4.6) \quad \|\varphi_1(t) - \varphi_2(t)\| \leq \|\varphi_1(t^*) - \varphi_2(t^*)\| e^{L|t-t^*|} (t, t^* \in [t_0, b]).$$

A 2.3. definíciót és a fenti egyenlőtlenséget felhasználva elsőként az $A_n = \|T_n(I(n)|t_n, h_n) - T(I(n)|t_n, h_n)\|$ mennyiséget becsüljük meg a következőképpen

$$\begin{aligned} A_n &= \left\| [y_n + h_n \psi(I(n)|t_n, y_n, Y_n(t_{n+1}), h_n) - Y_n(t_{n+1})] - \right. \\ &\quad \left. - [y(t_n) + h_n \psi(I(n)|t_n, y(t_n), y(t_{n+1}), h_n) - y(t_{n+1})] \right\| \leq \\ &\leq h_n K_{I(n)} (\|y_n - y(t_n)\| + \|Y_n(t_{n+1}) - y(t_{n+1})\|) + \\ &\quad + \|[y(t_{n+1}) - Y_n(t_{n+1})] - [y(t_n) - y_n]\| \leq \\ &\leq h_n K_{I(n)} (\|y_n - y(t_n)\| + \|y_n - y(t_n)\| e^{Lh_n}) + \\ &\quad + \left\| \int_{t_n}^{t_{n+1}} [f(t, y(t)) - f(t, Y_n(t))] dt \right\| \leq \\ &\leq h_n K_{I(n)} (1 + e^{Lh_n}) \|y_n - y(t_n)\| + L \int_{t_n}^{t_{n+1}} \|y(t) - Y_n(t)\| dt \leq \\ &\leq h_n [K_{I(n)} (1 + e^{Lh_n}) + L e^{Lh_n}] \|y_n - y(t_n)\|. \end{aligned}$$

Ha a (ψ, W, I) módszercsalád $t_n \in \Delta_N$ pontbeli globális hibáját $e_n = y_n - y(t_n)$ jelöli és teljesül, hogy $L \|\Delta_N\| \leq 1/2$, akkor

$$(4.7) \quad A_n \leq [K^* (1 + \sqrt{e}) + L \sqrt{e}] h_n \|e_n\| = Q h_n \|e_n\| \quad (n = 0, 1, \dots, N-1).$$

Ha $d_{n+1} = -T(I(n)|t_n, h_n)$ ($n=0, \dots, N-1$), akkor a (3.5) összefüggésből következik, hogy

$$\|e_{n+1}\| \leq (1 + h_n K_{I(n)}) \|e_n\| + \|T(I(n)|t_n, h_n)\| + h_n K_{I(n)} \|e_{n+1}\|,$$

ahonnan a (4.7) egyenlőtlenség miatt

$$\|e_{n+1}\| \leq [1 + h_n (K^* + Q)] \|e_n\| + \|T_n(I(n)|t_n, h_n)\| + h_n K^* \|e_{n+1}\|$$

következik. Ha feltesszük, hogy $\|\Delta_N\| K^* \leq 1/2$, akkor az

$$(1+a)/(1-b) \leq 1+2(a+b) \quad (a, b \geq 0, b \leq 1/2)$$

egyenlőtlenség alapján

$$\|e_{n+1}\| \leq [1 + 2h_n (2K^* + Q)] \|e_n\| + 2\|T_n(I(n)|t_n, h_n)\|.$$

Az $\|e_0\|=0$ feltétel miatt könnyen kimutatható (lásd [4]), hogy

$$(4.8) \quad \max_{0 \leq n \leq N} \|e_n\| \leq 2e^{2(b-t_0)(2K^*+Q)} \sum_{n=0}^{N-1} \|T_n(I(n)|t_n, h_n)\|$$

fennáll.

A $\|T_n(I(n)|t_n, h_n)\| \leq h_n \varepsilon$ ún. lépésegységenkénti hibafeltétel esetén a (4.8) egyenlőtlenségből következik, hogy

$$(4.9) \quad \max_{0 \leq n \leq N} \|e_n\| \leq c^* (b - t_0) \varepsilon = c_3 \varepsilon.$$

A második ún. lépésenkénti hibafeltétel esetén pedig (4.8)-ból

$$(4.10) \quad \max_{0 \leq n \leq N} \|e_n\| \leq c^* N \varepsilon = c_3(\Delta_N) \varepsilon$$

következik, amivel a tétel állításait teljes mértékben igazoltuk.

A lépésegységenkénti hibafeltétel használata esetén a c_3 konstans független a Δ_N felosztástól, míg a másik esetben $c_3 O(N)$ nagyságrendű. Ha a Δ_N felosztáshoz tartozó h_n ($n=0, 1, \dots, N-1$) lépéshosszakra teljesül, hogy $0 < Ah \leq h_n \leq Bh$ (A, B konstansok), akkor $\varepsilon = h^p$ ($p > 1$) esetén a (4.10) egyenlőtlenségből

$$(4.11) \quad \max_{0 \leq n \leq N} \|e_n\| \leq c_4 h^{p-1}$$

következik.

A gyakorlatban a (4.3)-beli c konstansokat a *Cauchy-problémától*, illetve a konkrét felosztástól függetlenül választják meg. A (4.3) feltétel ellenőrzését pedig $T_n(I(n)|t_n, h_n)$ valamely aszimptotikusan helyes becslése segítségével végzik. Mint-hogy ezek a becslések általában nem felső becslések, a (4.3), (4.4) feltételek velük végzett ellenőrzése nem garantálhatja a feltételek tényleges teljesülését. Ezért adott hibabecslő eljárások esetén vizsgálni kell a becslés pontosságát, azaz a becslés és $T_n(I(n)|t_n, h_n)$ eltéréseinek mértékét. Ha az eltérés $O(h_n \varepsilon)$, illetve $O(\varepsilon)$ nagyságrendű, akkor a fenti tétel bizonyítása alapján a lokális hibabecslési elv szerkezete és a rá vonatkozó tétel nem módosul.

A következőkben az egyik leggyakoribb hibabecslési eljárásra a lépésfelezésre vagy *Richardson-extrapolációra* igazoljuk a lokális hibabecslési elv aszimptotikus megfelelőjét.

Tegyük fel, hogy minden $f \in C^s$ és minden $w \in W$ esetén fennáll, hogy

$$(4.12) \quad T(w|t, h) = \Phi(w|t, y(t)) h^{p_w+1} + O(h^{p_w+2}),$$

ahol Φ az f -től és bizonyos rendű deriváltjaitól függ, $\Phi \in C^1$, az ordó tag pedig egyenletes az y függvényben és t -ben. Az s szám a (ψ, W, I) módszercsaládtól függ.

Legyen továbbá $\|h_n\| = \max_{0 \leq i \leq n} h_i$. Tegyük fel, hogy az y_n közelítést már kiszámoltuk és legyen $t_{n+1} = t_n + h$, $t_{n+2} = t_n + 2h$. Szintén tegyük fel, hogy az y_{n+1}, y_{n+2} közelítések kiszámítása során $I(n) = I(n+1) = \xi$ teljesül. Jelölje továbbá \bar{y}_{n+2} az $y(t_{n+2})$ azon közelítését, amelyet az

$$(4.13) \quad \bar{y}_{n+2} = y_n + 2h\psi(I(n)|t_n, y_n, \bar{y}_{n+2}, 2h)$$

formulával számítunk, azaz \bar{y}_{n+2} a (t_n, y_n) pontból $2h$ lépéshosszal számított közelítés.

Ekkor igaz a következő

4.2. TÉTEL. Ha a (ψ, W, I) egy lépéses módszercsalád indexfüggvénye kielégíti a $p_{I(i)} \geq p_\xi$ ($i=0, 1, \dots, n-1$) feltételt és a $\Delta_N \in \pi_\xi$ felosztásra $\|\Delta_N\| \leq h^*$ teljesül, akkor igaz, hogy

$$(4.14) \quad T_n(I(n)|t_n, h) = \frac{\bar{y}_{n+2} - y_{n+2}}{2^{p_\xi+1} - 2} + O(\|h_n\|^{p_\xi+2})$$

és

$$(4.15) \quad T_{n+1}(I(n)|t_{n+1}, h) = T_n(I(n)|t_n, h) + O(\|h_n\|^{p_\xi+2}).$$

Bizonyítás. Először azt igazoljuk, hogy

$$(4.16) \quad \mathbf{y}_{n+2} - \mathbf{Y}_n(t_{n+2}) = T_n(\xi|t_n, h) + T_n(\xi|t_{n+1}, h) + O(\|h_n\|^{p_\xi+2}).$$

Az egyenlőség bal oldala átalakítható a következőképpen

$$\begin{aligned} \mathbf{y}_{n+2} - \mathbf{Y}_n(t_{n+2}) &= [\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1})] + h[\psi(\xi|t_{n+1}, \mathbf{y}_{n+1}, \mathbf{y}_{n+2}, h) - \\ &\quad - \psi(\xi|t_{n+1}, \mathbf{Y}_n(t_{n+1}), \mathbf{Y}_n(t_{n+2}), h)] + \\ &\quad + [\mathbf{Y}_n(t_{n+1}) + h\psi(\xi|t_{n+1}, \mathbf{Y}_n(t_{n+1}), \mathbf{Y}_n(t_{n+2}), h) - \mathbf{Y}_n(t_{n+2})], \end{aligned}$$

ahol a harmadik szögletes zárójelben levő kifejezés definíció szerint $T_n(\xi|t_{n+1}, h)$. Vizsgáljuk most meg az első zárójelbeli kifejezés és $T_n(\xi|t_n, h)$ különbségét. A lokális hiba definíciója alapján

$$\begin{aligned} \|\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1}) - T_n(\xi|t_n, h)\| &= h\|\psi(\xi|t_n, \mathbf{y}_n, \mathbf{y}_{n+1}, h) - \\ &\quad - \psi(\xi|t_n, \mathbf{y}_n, \mathbf{Y}_n(t_{n+1}), h)\| \leq hK_\xi \|\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1})\|. \end{aligned}$$

Mint hogy az $\|A - B\| \leq \alpha \|A\|$ ($0 \leq \alpha < 1$) relációból $\|A\| \leq (1 - \alpha)^{-1} \|B\|$ és $\|A - B\| \leq \alpha(1 - \alpha)^{-1} \|B\|$ következik, azért a $\|A_N\| K^* \leq 1/2$ feltétel miatt igaz, hogy

$$\|\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1}) - T_n(\xi|t_n, h)\| \leq hK_\xi(1 - hK_\xi)^{-1} \|T_n(\xi|t_n, h)\| \leq 2hK_\xi \|T_n(\xi|t_n, h)\|.$$

A becslés jobb oldala a (4.7), (3.12) és (3.13) alapján

$$(4.17) \quad O(\|h_n\|^{\min\{p_{f(j)} | j=0, \dots, n\} + 2}) = O(\|h_n\|^{p_\xi+2})$$

nagyságrendű. A második szögletes zárójelben levő kifejezést a (2.5) feltétel, az előbbi okfejtés, a *Peano-egyenlőtlenség* és az $\|L\| \Delta_N \leq 1/2$ feltétel alapján felülről becsüli a

$$\begin{aligned} K_\xi(\|\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1})\| + \|\mathbf{y}_{n+2} - \mathbf{Y}_n(t_{n+2})\|) &\leq K_\xi(2\|T_n(\xi|t_n, h)\| + \\ &\quad + \|\mathbf{y}_{n+2} - \mathbf{Y}_{n+1}(t_{n+2})\| + \|\mathbf{Y}_{n+1}(t_{n+2}) - \mathbf{Y}_n(t_{n+2})\|) \leq K_\xi(2\|T_n(\xi|t_n, h)\| + \\ &\quad + 2\|T_{n+1}(\xi|t_{n+1}, h)\| + \|\mathbf{y}_{n+1} - \mathbf{Y}_n(t_{n+1})\| e^{Lh}) \leq 2K_\xi[(1 + \sqrt{e})\|T_n(\xi|t_n, h)\| + \\ &\quad + \|T_{n+1}(\xi|t_{n+1}, h)\|] \end{aligned}$$

menyiség. Emiatt a vizsgált egyenlőség jobb oldalának középső tagja szintén (4.17) nagyságrendű, amiből a (4.16) reláció fennállása már következik.

A lokális hiba (4.12) előállítására és $\Phi \in C^1$ alapján

$$\begin{aligned} \mathbf{y}_{n+2} - \mathbf{Y}_n(t_{n+2}) &= \Phi(\xi|t_n, \mathbf{Y}_n(t_n)) h^{p_\xi+1} + \Phi(\xi|t_{n+1}, \mathbf{Y}_n(t_{n+1})) h^{p_\xi+1} + \\ &\quad + O(\|h_n\|^{p_\xi+2}) = 2\Phi(\xi|t_n, \mathbf{y}_n) h^{p_\xi+1} + O(\|h_n\|^{p_\xi+2}). \end{aligned}$$

Mint hogy az eddigiek alapján

$$\begin{aligned} \bar{\mathbf{y}}_{n+2} - \mathbf{Y}_n(t_{n+2}) &= T_n(\xi|t_n, 2h) + O(\|h_n\|^{p_\xi+2}) = \\ &= \Phi(\xi|t_n, \mathbf{y}_n) (2h)^{p_\xi+1} + O(\|h_n\|^{p_\xi+2}), \end{aligned}$$

a két egyenlősből rendezéssel kapjuk, hogy

$$\Phi(\xi|t_n, y_n)h^{p_\xi+1} = \frac{\bar{y}_{n+2} - y_{n+2}}{2^{p_\xi+1} - 2} + O(\|h_n\|^{p_\xi+2}).$$

Innen (4.12) alapján a bizonyítandó (4.14) állítás már könnyen következik.

A (4.15) összefüggés helyessége a (4.12) előállításból és a $\Phi \in C^1$ feltételből következik.

Megjegyzések: 1. Explicit módszercsaládok esetén a tétel bizonyítása lényegesen egyszerűbb, ugyanis ekkor

$$y_{n+1} - Y_n(t_{n+1}) = T_n(\xi|t_n, h).$$

2. A hibabecslés pontossága, azaz a minimális $p_\xi+2$ rend, megfelel a korábban igazolt (3.8) és (3.13) kolátoknak. 3. A lépésfelezéses hibabecslés irodalomban ismert változataiban a nevező általában $2^{p_\xi}-1$. Ez a becslés azonban az

$$(4.18) \quad y_{n+2} - Y_n(t_{n+2})$$

hibára vonatkozik és nem a $T_n(\xi|t_n, h)$ hibára.

A szakaszban igazolt lokális hibabecslési elv alapján konkrét *Hull-típusú hatékonysági tételek* igazolhatók konkrét módszerekre (lásd [10]).

5. Kiegészítések és megjegyzések

Tetszőleges (ψ, W, I) egy lépéses módszercsalád konstrukciójánál a fő probléma a (ψ, W) bázismódszerek és az I indexfüggvény megválasztása. A (ψ, W) megválasztásával kapcsolatban számos érdekes, főleg heurisztikus eredmény ismeretes ([14], [15]). Az I indexfüggvény megválasztása hasonló lehet a DIFSUB eljárásához. Meg kell azonban jegyezni, hogy a *Richardson-extrapoláció* (és más hasonló hibabecslő eljárások) használata esetén az indexfüggvénynek olyannak kell lennie, hogy a formulák rendjére teljesüljön a 4.2. tételben megfogalmazott $p_{I(i)} \geq p_{I(i+1)}$ ($i \geq 0$) monotonitási feltétel. Ez a feltétel lényegében azért szükséges, mert az extrapoláció által kapott hibabecslés pontosságának összhangban kell lennie az eljárás tényleges globális hibájával.

A dolgozat eredményeit összehasonlítva a *Gear-típusú eljárásokra* nyert elméleti eredményekkel (lásd [7], [8], [11]) megállapítható, hogy az egy lépéses módszercsaládok sokkal rugalmasabbak a formula és a lépéshossz változtatást illetően. Az előbbiekben említett korlátozások is csak egy speciális hibabecslő eljárásra, illetve ennek használatára vonatkoznak. Amennyiben az egy lépéses módszercsalád olyan formulákból tevődik össze, amelyeknek egy pontra támaszkodó, a lépésfelezéssel azonos pontosságú hibabecsléseik vannak, a lépéshossz változtatására semmilyen korlátozás nem lép fel.

A szakirodalomban általánosan elterjedt az a vélemény, hogy a *Runge—Kutta-módszerek*, főleg az implícitek, sokkal gépidőigényesebbek, mint például a lineáris több lépéses módszerek. Azonban GUPTA [9] dolgozata szerint ez nincs mindig így, sőt bizonyos GUPTA által tesztelt problémaosztályokon az ellenkezője igaz. Ez azt jelenti, hogy adott problémaosztályokon az egy lépéses módszercsaládok a gépidő vonatkozásában is versenyképesebbek lehetnek, mint a *Gear-típusú eljárások*.

Végül vizsgáljuk meg végtelen sok formulából álló egylépéses módszercsaládok konvergenciájának lehetőségét. Ezt a vizsgálatot az indokolja, hogy WERSCHULZ [18] dolgozatában, az egylépéses módszerek komplexitását vizsgálva, igazolta, hogy az egylépéses módszerek optimális rendje tart a végtelenhez, hacsak az előírt $\varepsilon^* > 0$ hibakorlát tart a zérushoz. Ez azt jelenti, hogy az olyan módszercsaládok, amelyekben minden ε^* hibakorláthoz van optimális rendű (azaz minimális komplexitású) módszer, szükségképpen végtelen sok formulából kell hogy álljanak. A következőkben igazoljuk, hogy léteznek olyan végtelen sok formulából álló egylépéses módszercsaládok, amelyek adott problémaosztály esetén konvergensek.

Tekintsük az

$$(5.1) \quad y' = Ay + F(t); \quad y(t_0) = y_0 \quad (t_0 \geq 0)$$

alakú *lineáris Cauchy-problémákat*, ahol A $m \times m$ típusú konstans mátrix. Ezek megoldására igaz, hogy

$$y_n - y(t_n) = \sum_{j=1}^n [Y_j(t_n) - Y_{j-1}(t_n)] = \sum_{j=1}^n [y_j - Y_{j-1}(t_j)] \exp(A(t_n - t_j)).$$

Ha az A mátrix sajátértékei a $\{\operatorname{Re} z < 0\}$ félsíkba esnek, akkor létezik két $u_1, u_2 > 0$ konstans, hogy $\|\exp(At)\| \leq u_1 e^{-u_2 t}$ ($t \geq 0$). Ebből következik, hogy

$$(5.2) \quad \|y_n - y(t_n)\| \leq u_1 \sum_{j=1}^n \|y_j - Y_{j-1}(t_j)\| e^{-u_2(t_n - t_j)}.$$

Tegyük most fel, hogy a (\tilde{y}, \tilde{W}) módszerhalmaz az olyan

$$(5.3) \quad y_{n+1} - y_n = h_n \sum_{i=1}^s c_i k_i$$

$$k_i = f\left(t_n + a_i h_n, y_n + h_n \sum_{j=1}^s b_{ij} k_j\right) \quad (i = 1, \dots, s)$$

alakú *Runge—Kutta-módszerekből* áll, amelyek a következő tulajdonsággal rendelkeznek: Ha a módszert konstans $h > 0$ lépéshosszal az

$$(5.4) \quad y' = \lambda y; \quad y(0) = 1 \quad (\lambda \in \mathbb{C})$$

modellegenletre alkalmazzuk, akkor az első lépés eredményeként az $y_1 = r_{k,s}(z)$ közelítés adódik, ahol $r_{k,s}(z)$ az e^z ($z \in \mathbb{C}$) függvény (k, s) -indexű *Padé-approximációja*, $k \in \{s, s+1, s+2\}$, $s \geq 1$ és $z = \lambda h$.

Az $r_{k,s}(z)$ *Padé-approximációt* az

$$(5.5) \quad r_{k,s}(z) = N_{k,s}(z)/D_{k,s}(z)$$

$$N_{k,s}(z) = \sum_{m=0}^s \frac{(k+s-m)! s!}{(k+s)! m! (s-m)!} z^m$$

$$D_{k,s}(z) = \sum_{m=0}^k \frac{(k+s-m)! k!}{(k+s)! m! (k-m)!} (-z)^m$$

összefüggések definiálják. Ismeretes ([3]), hogy minden $s \geq 1$, $k \in \{s, s+1, s+2\}$ esetén léteznek a fenti tulajdonsággal rendelkező $(k+s)$ -ed rendű *implicit Runge—Kutta-módszerek*. Ezért a (\tilde{y}, \tilde{W}) módszerhalmaz végtelen sok módszerből áll.

Legyen a továbbiakban $F=0$ és tegyük fel, hogy az

$$\tilde{I} = \tilde{I}(\Delta_N): \{0, 1, \dots, N\} \rightarrow \mathbb{N}$$

indexfüggvény az n -edik lépésben az $r_{k(n), s(n)}(z)$ *Padé-approximációt* generáló *Runge—Kutta-formulát* választja ki, ahol $k(n) \in \{s(n), s(n)+1, s(n)+2\}$, $s(n) \geq 1$, $n \geq 0$. Ekkor

$$y_{j+1} = r_{k(j), s(j)}(Ah_j)y_j \quad (j = 0, 1, \dots, N-1)$$

és

$$\|y_n - y(t_n)\| \leq u_1 \sum_{j=0}^{n-1} \|\exp(Ah_j) - r_{k(j), s(j)}(Ah_j)\| \|y_j\|.$$

Ha A hermitikus és negatív definit, akkor spektrál normában

$$(5.6) \quad \|\exp(Ah) - r_{k,s}(Ah)\| = \max_{1 \leq i \leq m} |e^{\lambda_i h} - r_{k,s}(\lambda_i h)|$$

és

$$(5.7) \quad \|r_{k,s}(Ah)\| = \max_{1 \leq i \leq m} |r_{k,s}(\lambda_i h)|,$$

ahol λ_i az A mátrix i -edik sajátértéke. EHLE ([1]) igazolta, hogy minden $s \geq 1$, $k \in \{s, s+1, s+2\}$ esetén fennáll az

$$(5.8) \quad |r_{k,s}(z)| \leq 1 \quad (\operatorname{Re} z < 0)$$

egyenlőtlenség, amiből (5.7) miatt $\|r_{k,s}(Ah)\| \leq 1$ ($h > 0$) következik. ELLIOTT [2] dolgozata alapján az $r_{k,s}(z)$ *Padé-approximáció* hibájára fennáll, hogy

$$(5.9) \quad D_{k,s}(z)e^z - N_{k,s}(z) = \frac{z^{k+s+1}}{2\pi i} \int_G \frac{e^t D_{k,s}(t)}{(t-z)^{k+s+1}} dt,$$

ahol G a $0, z \in \mathbb{C}$ pontokat a belsejében tartalmazó zárt görbe. Az $r_{k,s}(z)$ törtfüggvény (5.5) előállítására alapján

$$|D_{k,s}(z)| > 2/3 \quad (|z| \leq 1/4), \quad |N_{k,s}(z)| < 2^{s+1} \quad (|z| \leq 2).$$

Legyen a továbbiakban $|z| \leq 1/4$ és G a $\{t \in \mathbb{C} \mid |t|=2\}$ körvonallal azonos. Ekkor $|t-z| > 3/2$ és

$$|e^z - r_{k,s}(z)| \leq |D_{k,s}^{-1}(z)| \frac{|z^{k+s+1}|}{2\pi} \int_G \frac{|e^t| |D_{k,s}(t)|}{|t-z|^{k+s+1}} dt \leq 80 \left(\frac{8}{9}\right)^s |z|^{k+s+1}.$$

Ebből következik, hogy

$$\|y_n - y(t_n)\| \leq 80u_1 \|y_0\| \sum_{j=0}^{n-1} |\lambda_{i(j)} h_j|^{(k(j)+s(j)+1)},$$

ahol az $i(j)$ indexet az

$$\|\exp(Ah_j) - r_{k(j), s(j)}(Ah_j)\| = |e^{\lambda_{i(j)} h_j} - r_{k(j), s(j)}(\lambda_{i(j)} h_j)|$$

egyenlőség definiálja. Feltéve, hogy $|\lambda_i| \sqrt{\|\Delta_N\|} \leq 1$ ($i = 1, \dots, m$) kapjuk, hogy

$$(5.10) \quad \|y_n - y(t_n)\| \leq Q_1 \sum_{j=0}^{n-1} h_j^{k(j)+s(j)+1/2} \leq Q_1(b-t_0) \sqrt{\|\Delta_N\|}.$$

Az (5.10) alapján azonnal következik az

5.1. TÉTEL. A végtelen $(\tilde{\psi}, \tilde{W}, \tilde{I})$ módszercsalád tetszőleges nem korlátos \tilde{I} indexfüggvény esetén konvergens az

$$(5.11) \quad y' = Ay; \quad y(t_0) = y_0 \quad (t_0 \geq 0)$$

alakú differenciálegyenletek osztályán, ahol A $m \times m$ típusú konstans, negatív definit és hermitikus mátrix.

IRODALOM

- [1] EHLE, B. L., "A-stable methods and Padé approximations to the exponential", *SIAM J. Math. Anal.* **4** (1973) 671—680.
- [2] ELLIOTT, D., "Truncation errors in Padé approximations to certain functions: an alternative approach", *Math. Comp.* **21** (1967) 398—406.
- [3] FUCHS, F., "A-stability of Runge—Kutta methods with single and multiple nodes", *Computing* **16** (1976) 39—48.
- [4] GALÁNTAI, A., "Convergence theorems and error analysis for one-step methods", *Annales Univ. Sci. Budapest, Sect. Math.* **19** (1976) 69—78.
- [5] GALÁNTAI, A., „Vizsgálatok a közönséges differenciálegyenletek közelítő módszereinek konvergencia és hibaanalízisének körében”, kandidátusi értekezés, Budapest, 1978.
- [6] GEAR, C. W., *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971).
- [7] GEAR, C. W. and TU, K. W., "The effect of variable meshsize on the stability of multistep methods", *SIAM J. Numer. Anal.* **11** (1974) 1025—1043.
- [8] GEAR, C. W. and WATANABE, D. S., "Stability and convergence of variable multistep methods", *SIAM J. Numer. Anal.* **11** (1974) 1044—1058.
- [9] GUPTA, G. K., "A note about overhead costs in ODE solvers", *ACM Transactions on Math. Software* **6** (1980) 319—326.
- [10] HULL, T. E., "The numerical integration of ordinary differential equations", IFIP 68, North-Holland, Amsterdam, (1969) 40—53.
- [11] MÁRZ, R., "Variable multistep methods", Seminarbericht 7, Sektion Mathematik der Humboldt-Universität zu Berlin, 1979.
- [12] PLANT, R. E., "The efficient numerical solution of biological simulation problems", *Computer Programs in Biomedicine* **10** (1979) 1—15.
- [13] SCHMIDT, J. W., „Über lineare Ungleichungen vom Gronwall'schen Typ“, *Beiträge zur Num. Math.* **5** (1976) 171—189.
- [14] SHAMPINE, L. F., WATTS, H. A., "Practical solution of ordinary differential equations by Runge—Kutta methods", Report, SAND 76-0585, Sandia Laboratories, December, 1976.
- [15] SHAMPINE, L. F. and WISNIEWSKI, J. A., "The variable order Runge—Kutta code RKSU and its performance", Report, SAND 78-1347, Sandia Laboratories, December, 1978.
- [16] STETTER, H. J., *Analysis of Discretization Methods for Ordinary Differential Equations* (Springer, Berlin, 1973).
- [17] VARGA, R. S., *Functional Analysis and Approximation Theory in Numerical Analysis* (SIAM, Philadelphia, 1971).
- [18] WERSCHULZ, A. G., "Optimal order and minimal complexity of one-step methods for initial value problems", Report, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, September, 1976.
- [19] ZLATEV, Z., "Stability properties of variable stepsize variable formula methods", *Numer. Math.* **31** (1978) 175—182.

(Beérkezett: 1982. november 25.)

GALÁNTAI AURÉL

AGRÁRTUDOMÁNYI EGYETEM MATEMATIKAI ÉS SZÁMÍTÁSTECHNIKAI INTÉZET
2103 GÖDÖLLŐ

CONVERGENCE THEOREMS AND ERROR ANALYSIS
FOR THE FAMILIES OF ONE-STEP METHODS

A. GALÁNTAI

In this paper we investigate the convergence and error analysis of the families of one-step methods defined by (2.3) and (2.6). The stability is proven by the two-sided error estimation (3.2), and this result is used to prove the convergence of such processes in the sense (2.7) and to estimate its speed [see (3.8) and (3.13)]. The principle of local error estimation, that is condition (4.4) implies (4.5), is proven. We prove that *Richardson-extrapolation* asymptotically satisfies this principle. The convergence of infinite families of *Runge—Kutta methods* based on *Padé approximations* $r_{k,s}(z)$ to e^z ($k \in \{s, s+1, s+2\}$, $s \geq 1$) is also proven for the class of linear differential equations (5.11) with negative definite and hermitian coefficient matrix.

A PERTURBÁLT VAN DER POL-DIFFERENCIÁLEGYENLET SZÁMÍTÓGÉPES VIZSGÁLATA

SPARING LÁSZLÓ

Budapest

A nem lineáris rezgések elméletének egyik fontos, sokszor használt példája a *van der Pol-egyenlet*. Ebben a dolgozatban kvalitatív eredményekre támaszkodva, számítógépes módszerrel numerikus eredményeket adunk és rajzokkal illusztráljuk azokat. A kidolgozott programok segítségével irányítottan periodikus perturbációkra az aszimptotikusan stabilis periodikus megoldás meghatározható és szemléltethető.

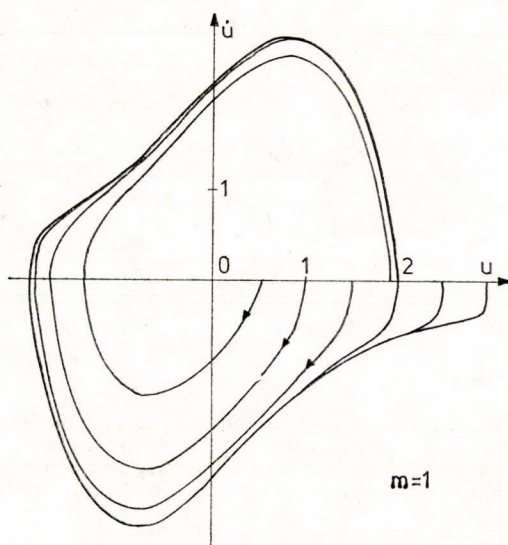
1. Az autonóm egyenlet

Az $\ddot{u} + m(u^2 - 1)\dot{u} + u = 0$, $u: R \rightarrow R$, m pozitív állandó *van der Pol-differenciálegyenlet* a nem lineáris rezgések elméletének egyik alapegyenlete. A pont a t idő szerinti deriválást jelenti. Jó közelítéssel írja le például egy elektroncsöves visszacsatolt oszcillátor rezgéseit [1]. A közönséges differenciálegyenletek elméletének is egyik fontos mintapéldája. Ismeretes, hogy létezik (a t független változó eltolásaitól eltekintve) pontosan egy nem állandó periodikus megoldása, amely aszimptotikusan orbitálisan stabilis. Ennek a periodikus u_0 megoldásnak a kezdeti értékét megadhatjuk

$$u_0(0) = a > 0, \quad \dot{u}_0(0) = 0$$

formában. Legyen u_0 legkisebb pozitív periódusa τ_0 , vagyis $u_0(t + \tau_0) = u_0(t)$ minden valós t -re. Az a és τ_0 értékek az m paraméter függvényei, az irodalomban több közelítő módszer található meghatározásukra, pl. [2] és [5]. Ezek a módszerek azonban meglehetősen sok kézi számolást igényelnek. Számítógép alkalmazásával az orbitális stabilitást nagyon jól fel lehet használni. Tetszőleges kezdeti értékből indítva a megoldást, az elmélet szerint a trajektória a határciklushoz tart. Tudjuk, hogy $\lim_{m \rightarrow 0} a = 2$, így nagyságrendileg ismert a értéke. A trajektóriákat rajzoló program eredményei szerint a határciklus rövid idő alatt magához vonzza a megoldásokat.

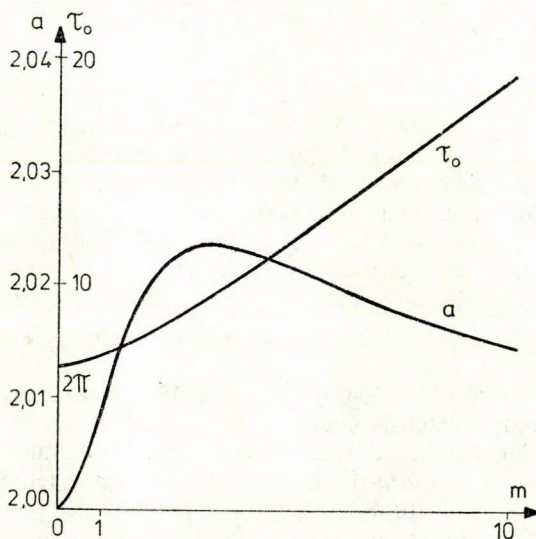
A számításokat ODRA—1204 típusú számítógépen végeztük, az eredményeket CIL típusú plotteren jelentettük meg. A programok ALGOL nyelven készültek. A differenciálegyenlet megoldására egy, a számítógép programkönyvtárban található negyedrendű *Runge—Kutta-típusú eljárást* használtunk fel, amelyet erre a célra átalakítottunk és felgyorsítottunk. Az eljárás egyes paramétereivel a számítási hiba korlátozható. Ellenőrzésül az $m=0$ eset kínálkozik, amikor a harmonikus oszcillátor megoldása pontosan ismert. A kör alakú trajektóriánál a sugár relatív hibája



1. ábra

a kiválasztott korlátozó paramétereknél egy körülfordulás után $2 \cdot 10^{-7}$, a periódus relatív hibája $2 \cdot 10^{-6}$ volt. Így a és τ_0 meghatározásának relatív hibája kisebb 10^{-5} -nél.

Az 1. táblázat tartalmazza a numerikus eredményeket a -ra és τ_0 -ra néhány m érték esetén.



2. ábra

1. TÁBLÁZAT

m	a	τ_0
0,1	2,000 059	6,287 112
0,2	2,000 391	6,298 876
0,3	2,000 902	6,318 443
0,4	2,001 595	6,345 747
0,5	2,002 480	6,380 676
0,6	2,003 488	6,423 103
0,7	2,004 640	6,472 835
0,8	2,005 912	6,529 638
0,9	2,007 223	6,593 237
1,0	2,008 598	6,663 290
1,1	2,010 005	6,739 423
1,2	2,011 377	6,821 229
1,3	2,012 717	6,908 283
1,4	2,014 015	7,000 139
1,5	2,015 203	7,096 376
1,6	2,016 328	7,196 577
1,7	2,017 377	7,300 355
1,8	2,018 292	7,407 366
1,9	2,019 131	7,517 307
2,0	2,019 890	7,629 874
2,5	2,022 330	8,224 723
3,0	2,023 295	8,859 093
3,227	2,023 400	9,156 120
3,5	2,023 362	9,521 080
4,0	2,022 950	10,204 526
4,5	2,022 288	10,901 765
5,0	2,021 495	11,612 263
6,0	2,019 816	13,061 881
7,0	2,018 206	14,539 684
8,0	2,016 729	16,038 147
9,0	2,015 427	17,552 222
10,0	2,014 270	19,078 361

A későbbi vizsgálatoknál szükség lesz az autonóm egyenlettel kapcsolatos néhány további mennyiségre. A

$$(1.1) \quad \ddot{v} + m[u_0^2(t) - 1]\dot{v} + v = 0$$

homogén lineáris differenciálegyenlet egyik megoldása nyilván az $u_0(t)$ függvény. Jelölje v_0 (1.1) azon megoldását, amelyre

$$v_0(0) = 0, \quad \dot{v}_0(0) = \frac{1}{a}.$$

A *van der Pol*-egyenlettel ekvivalens rendszer megfelelő variációs rendszerének karakterisztikus multiplikátorai 1 és $W(\tau_0)$, ahol

$$W(r) = e^{-m \int_0^r [u_0^2(t) - 1] dt}.$$

2. Irányítottan periodikus perturbációk

A gyakorlatban megvalósuló rezgő rendszerekre mindig hatnak zavarások, azonban egyes esetekben ezek frekvenciáját tudjuk — bizonyos határokon belül — szabályozni. Ezért lényeges az irányítottan periodikus perturbációk matematikai vizsgálata. Ezeknél a zavarásoknál a frekvenciát független változónak tekintjük.

Vizsgáljuk a *van der Pol-egyenletet* a következő irányítottan periodikus perturbációval:

$$(2.1) \quad \ddot{u} + m(u^2 - 1)\dot{u} + u = \mu\gamma\left(\frac{t}{\tau} u, \dot{u}, \mu, \tau\right),$$

ahol $\mu \in R$, γ analitikus függvénye a változónak és minden rögzített τ -ra a t változóban periodikus τ periódussal. Jelöljük (2.1)

$$(2.2) \quad \begin{aligned} u(\vartheta; \vartheta, a+h, \mu, \tau) &= a+h \\ u(\vartheta; \vartheta, a+h, \mu, \tau) &= hm(1-a^2-ah-h^3/3) \end{aligned}$$

kezdeti értéknek eleget tevő megoldását $u(t; \vartheta, a+h, \mu, \tau)$ -val. Fennáll a következő

2.1. TÉTEL: Ha $W(\tau_0) < 1$, akkor minden (μ, ϑ) -ra, melyre $|\mu|$ és $|\vartheta|$ elég kicsi,

a) létezik pontosan egy $\tau(\mu, \vartheta)$ periódus és pontosan egy $h(\mu, \vartheta)$ szám, hogy az

$$(2.3) \quad \ddot{u} + m(u^2 - 1)\dot{u} + u = \mu\gamma\left(\frac{t}{\tau(\mu, \vartheta)}, u, \dot{u}, \mu, \tau(\mu, \vartheta)\right)$$

egyenlet $u(t; \vartheta, a+h(\mu, \vartheta), \mu, \tau(\mu, \vartheta))$ megoldása periodikus $\tau(\mu, \vartheta)$ periódussal, a τ és h függvény a $(\mu, \vartheta) = (0, 0)$ pont egy környezetében analitikus, $\tau(0, 0) = \tau_0$, $h(0, 0) = 0$ és $u(t; 0, a, 0, \tau_0) = u_0(t)$.

b)

$$(2.4) \quad \tau(\mu, \vartheta) = \tau_0 + \mu\tau_1(\vartheta) + o(\mu)$$

ahol

$$\tau_1(\vartheta) = \int_0^{\tau_0} \frac{\gamma^*(r)}{W(r)} \left[\dot{v}_0(r) - \frac{v_0(\tau_0)}{a(1-W(\tau_0))} \dot{u}_0(r) \right] dr,$$

$$\gamma^*(t) = \gamma\left(\frac{t+\vartheta}{\tau_0}, u_0(t), \dot{u}_0(t), 0, \tau_0\right), \quad t \in R.$$

c) A $(\mu, \vartheta) = (0, 0)$ pontnak van olyan U környezete, melyre ha $(\mu, \vartheta) \in U$ és $\mu \cdot \lambda'_\mu(0, 0) < 0$, ahol

$$\lambda'_\mu(0, 0) = \int_0^{\tau_0} \frac{1}{W(t)} \cdot \gamma'_t\left(\frac{t}{\tau_0} u_0(t), \dot{u}_0(t), 0, \tau_0\right) \left[\dot{v}_0(t) - \frac{v_0(\tau_0)\dot{u}_0(t)}{a(1-W(\tau_0))} \right] dt,$$

akkor a (2.3) egyenlet $u(t; \vartheta, a+h(\mu, \vartheta), \mu, \tau(\mu, \vartheta))$ megoldása aszimptotikusan stabilis. [3], [4].

Ez a tétel biztosítja, hogy kis $|\mu|$ és $|\vartheta|$ esetén a perturbált egyenletnek is van periodikus megoldása, és megfelelő előjelű μ esetén a feltételek fennállásakor ez a megoldás aszimptotikusan stabilis.

A számítógépes vizsgálatok során a (2.1) egyenlet m paraméterének néhány értékére ellenőriztük a 2.1. tétel feltételeit, majd a képletekkel megadott $\lambda'_\mu(0, 0)$ és $\tau_1(9)$ megfelelő pontosságú meghatározásával lehetővé vált az aszimptotikusan stabilis megoldások megkeresése.

2. TÁBLÁZAT

m	$W(\tau_0)$	$\lambda'_\mu(0, 0)$	$\tau_1(0)$
0,1	0,533	1,574	0,027
0,2	0,283	1,581	0,042
0,3	0,149	1,593	0,064
0,4	0,0770	1,610	0,077
0,5	0,0392	1,627	0,122
0,6	0,0195	1,649	0,153
0,7	0,00940	1,674	0,191
0,8	0,00434	1,702	0,233
0,9	0,00198	1,728	0,284
1,0	0,000860	1,756	0,340
1,1	0,000357	1,776	0,402
1,2	0,000141	1,837	0,457
1,3	0,0000532	1,827	0,539
1,4	0,0000191	2,146	0,493
1,5	0,00000647	2,087	0,626
1,6	0,00000208	2,520	0,537
1,7	0,000000633	9,240	-1,883
1,8	0,000000182	6,594	-1,574
1,9	0,0000000496	19,143	-8,673
2,0	0,0000000127	-4,570	-42,165

A következő eredmények a $\gamma\left(\frac{t}{\tau}\right) = \sin \frac{2\pi t}{\tau}$ függvényre vonatkoznak, de a program alkalmas tetszőleges perturbáció kezelésére. Adott μ és 9 mellett a $\tau \approx \tau_0 + \mu\tau_1(9)$ közelítés felhasználásával gyökközelítő módszerek segítségével (2.2) alapján a $h(\mu, 9)$ függvény értékét is néhány jegyre meghatároztuk.

3. Eredmények

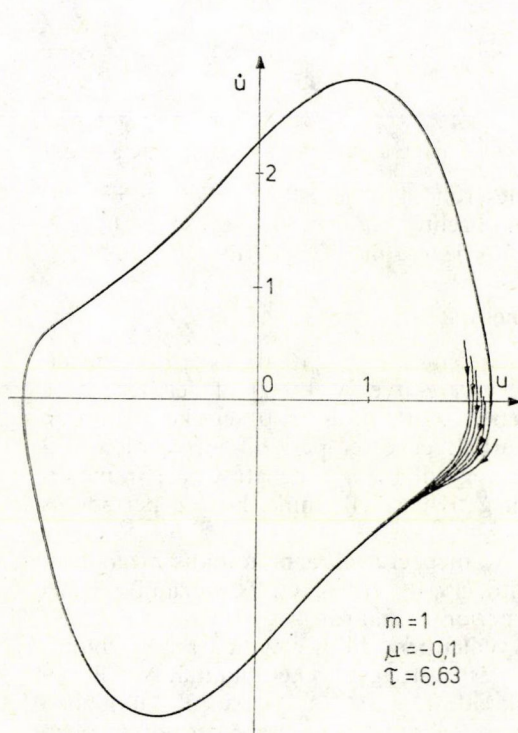
A 2. táblázat tartalmazza néhány m esetén $W(\tau_0)$, $\lambda'_\mu(0, 0)$ és $\tau_1(0)$ értékét, a becsült számítási pontosságnak megfelelően kerekítve. A $W(\tau_0) < 1$ feltétel ezekre az m -ekre teljesül, azonban $m=2$ közelében $\lambda'_\mu(0, 0)$ és $\tau_1(0)$ értéke jelentősen változik és előjelet is vált, így itt a perturbált egyenlet periodikus megoldásával problémák lehetnek. Az is megállapítható, hogy pl. $m=1$ mellett a μ paraméternek a választott szinuszos zavarás esetén negatívnak kell lennie, hogy a periodikus megoldás aszimptotikusan stabilis legyen.

Kiragadva az $m=1$ esetet, néhány μ -re megkerestük a periodikus megoldást. A 3. táblázat tartalmazza a kapott közelítő $h(\mu, 0)$ értékeket, és a számítógéppel generált, illetve a (2.4) formulával kapott periódus arányát.

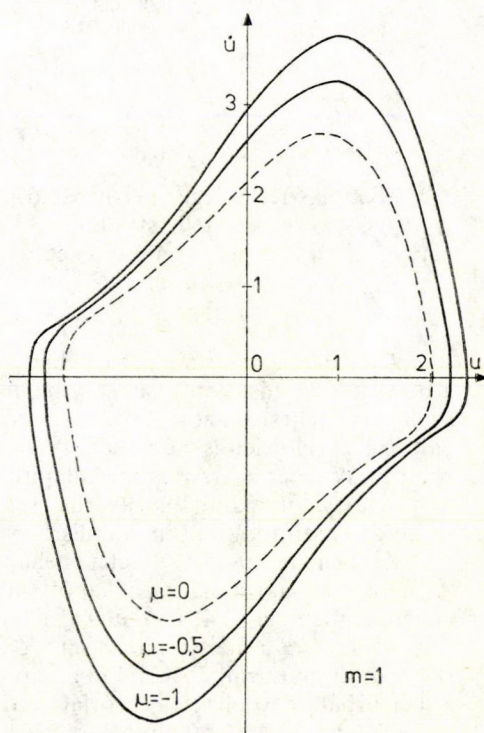
Az 1. ábra a fázissíkon az autonóm egyenlet periodikus megoldásának orbitális stabilitását illusztrálja. A 2. ábra a kapott a és τ_0 függvényeket mutatja. A 3. ábrán a perturbált egyenlet trajektóriái láthatók adott m, μ és τ esetén, különböző h értékekre. A 4. ábra adott m -re különböző μ értékekhez tartozó periodikus megoldások trajektóriáit mutatja.

3. TÁBLÁZAT

μ	$h(\mu, 0)$	τ	$\tau_0 + \mu\tau_1(0)$	$\frac{\tau}{\tau_0 + \mu\tau_1(0)}$
0,2	-0,121	6,792	6,731	1,009
0,1	-0,054	6,711	6,697	1,002
0,0	0,0	6,663	6,663	1,0
-0,1	0,041	6,639	6,629	1,001
-0,2	0,072	6,630	6,595	1,005
-0,3	0,095	6,631	6,561	1,011
-0,4	0,112	6,638	6,527	1,017
-0,5	0,125	6,649	6,493	1,024
-0,6	0,136	6,658	6,459	1,031
-0,7	0,144	6,672	6,425	1,038
-0,8	0,150	6,686	6,392	1,046
-0,9	0,156	6,699	6,358	1,054
-1,0	0,159	6,716	6,324	1,062



3. ábra



4. ábra

IRODALOM

- [1] CESARI, L., *Asymptotic behavior and stability problems in ordinary differential equations* (Springer, Berlin, 1963).
- [2] DAVIS, R. T. and ALFRIEND, K. T., "Solution to van der Pol's equation using a perturbation method" *Int. J. Non-Linear Mechanics* 2 (1967) 153—162.
- [3] FARKAS, I. and FARKAS, M., "On perturbations of van der Pol's equation" *Annales Univ. Sci. Budapest, Sectio Math.* 15 (1972) 155—164.
- [4] FARKAS, M., "Autonóm rendszerek periodikus perturbációiról", *Alkalmazott Matematikai Lapok* 1 (1975) 197—254.
- [5] URABE, M., *Nonlinear Autonomous Oscillations* (Academic Press, New York, 1967).

(Beérkezett: 1982. július 23.)

SPARING LÁSZLÓ

BME GÉPÉSZMÉRNÖKI KAR MATEMATIKAI TANSZÉK
1521 BUDAPEST, STOCZEK U. H. ÉP. IV. EM. 41.

NUMERICAL INVESTIGATION OF THE PERTURBED
VAN DER POL'S EQUATION

L. SPARING

Numerical results are given concerning the periodically perturbed van der Pol's equation. Applying the theory of M. FARKAS we give numerical conditions for the stability of the perturbed periodic solutions. The solutions themselves and their periods are determined. We illustrate the attractivity of these solutions by computer patterns.

EREDETI CSÚCSPONT KERESÉSE ÉS ALKALMAZÁSA KONKÁV FÜGGVÉNY LINEÁRIS FELTÉTELEK MELLETTI MINIMALIZÁLÁSÁKOR

FÜLÖP JÁNOS

Budapest

Konkáv függvényt minimalizálunk csúcsponttal rendelkező konvex poliéderen. Ha a függvény alulról korlátos a poliéderen, akkor a feladatnak van véges optimuma és azt a poliéder egy csúcspontján is felveszi. Ezen feladat megoldására véges metszősík módszert nyerünk konvex metszések eredeti csúcspontból történő előállításával. A dolgozat az eredeti csúcspontot kereső *Majthay—Whinston-módszer* egy módosítását mutatja be és alkalmazza a konkáv minimalizálási feladat megoldásánál.

1. Bevezetés

Tekintsük a következő matematikai programozási feladatot:

$$(1.1) \quad \begin{aligned} &\min f(\mathbf{x}) \\ &\mathbf{Ax} = \mathbf{b} \\ &\mathbf{x} \geq \mathbf{0}, \end{aligned}$$

ahol \mathbf{A} $m \times n$ -es mátrix, $\mathbf{x} \in R^n$, $\mathbf{b} \in R^m$, $\text{rang}(\mathbf{A}) = m$ és a

$$P_0 = \{\mathbf{x} | \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

poliéder nem üres. Az $f(\mathbf{x})$ függvényről feltesszük, hogy R^n -ben konkáv és a P_0 poliéderen alulról korlátos. Ekkor az (1.1) feladatnak van véges optimuma és az a P_0 egy csúcspontján is felvétetik [12]. Ugyanez teljesül akkor is, ha $f(\mathbf{x})$ kvázi-konkáv és P_0 korlátos [9]. Tetszőleges olyan feladat, ahol egy konkáv függvényt minimalizálunk csúcsponttal rendelkező poliéderen (1.1) alakra transzformálható és megtehető ez a kvázikonkáv esetben is [2].

Az (1.1) feladat egy gyakorlati alkalmazásaként említhető a termeléstervezési feladat csökkenő ütemben növekvő költség minimalizálása esetén [13]. Számos speciális nemkonvex programozási feladat írható fel vagy transzformálható (1.1) alakra, például a komplementaritási feladat, a bilineáris feladat stb.

Az (1.1) feladat egy megoldási módja lehetne a P_0 poliéder csúcspontjainak a leszámllálása. Léteznek implicit leszámllálási módszerek, melyek esetleg felhasználják a célfüggvény valamely lineáris alsó becslését is [1, 2, 10], azonban jelentős tár- és időigényük miatt nagyobb méretű feladatok megoldására nem tűnnek felhasználhatónak.

Az (1.1) feladat megoldását célzó módszerek egy másik csoportját alkotják a metszősík módszerek. Ezek a következőképpen működnek. Kiválasztjuk a P_0

poliéder egy x^0 csúcspontját és előállítunk egy olyan $t'x = t_0$ hipersíkot, amelyre $t'x^0 > t_0$, valamint tetszőleges

$$x \in P_0 \cap \{x | t'x > t_0\}$$

esetén $f(x) \cong f^0$, ahol f^0 az addig talált legjobb célfüggvényérték. A

$$(1.2) \quad t'x \leq t_0$$

feltételt a P_0 -t meghatározó feltételekhez csatoljuk és az így nyert, P_1 -gyel jelölt feltételi poliéderen folytatjuk a metszősík eljárást. Általában az eljárás minden fölépésében valamely P_k poliéderből egy szűkebb P_{k+1} poliédert állítunk elő egy újabb feltétel csatolásával. A metszősík módszer akkor fejeződik be, ha a megengedett halmaz üressé válik. A metszősík előállítására általában a konvex és polár metszéseket [2] használják. Ennek az alapmódszernek a végeessége vagy konvergenciája azonban még korlátos P_0 esetén sem biztosítható.

Tudjuk, hogy az (1.1) feladat optimumát elég P_0 csúcspontjain keresni. Tegyük fel, hogy adott egy eljárás, amely megkeresi az aktuális P_k poliédernek egy olyan csúcspontját, amely csúcspontja P_0 -nak is, illetve jelzi, ha már ilyen nincsen. Az ilyen csúcspontokat eredeti csúcspontoknak nevezzük. Ha az (1.2) metszést egy x^0 eredeti csúcspontból állítjuk elő, akkor x^0 már nem lesz benne az (1.2) csatolásával nyert P_{k+1} poliéderben. Mivel P_0 -nak véges számú csúcspontja van, ezért véges számú (1.2) metszés után az aktuális poliéder már nem tartalmaz eredeti csúcspontot. Ekkor befejezhetjük a metszősík módszert, mivel tudjuk, hogy a célfüggvény nem vesz fel a maradék poliéderen az eddig talált legjobbnál jobb értéket.

Az eredeti csúcspontot kereső, MAJTHAY és WHINSTON által adott eljárás [9] minden fölépésben vagy egy eredeti csúcspontot talál, vagy jelzi, hogy az aktuális poliéderben nincs eredeti csúcspont, vagy a P_0 poliédernek egy olyan határoló felületét állítja elő, amelynek az aktuális poliéderbe eső része nem üres, de ezen a részen már nincs eredeti csúcspont. Ez utóbbi esetben az aktuális poliéderről levágja a P_0 -nak ezt az eredeti csúcspontot nem tartalmazó határoló felületét, anélkül, hogy ezzel bármely, az aktuális poliéderbe eső eredeti csúcspontot kizárna. Mivel P_0 -nak véges sok határoló felülete van és P_0 minden csúcspontja egyben határoló felülete is P_0 -nak, ezért véges számú fölépésben egy eredeti csúcsponthoz jutunk, vagy megbizonyosodunk arról, hogy nincs már eredeti csúcspont.

A Majthay—Whinston-módszernek itt bemutatandó módosítása lehetővé teszi a [9]-ben előállított extrémális határoló felületet tartalmazó nagyobb dimenziós felületeken annak kimutatását, hogy ott már nincs eredeti csúcspont. Ez egyben az eredeti csúcspontot kereső eljárás, így az (1.1) feladatot megoldó módszer gyorsítását is jelenti.

A 2. fejezet egy, a Majthay—Whinston-módszer közvetlen módosításaként értelmezhető eljárást mutat be eredeti csúcspontot már nem tartalmazó határoló felületek megkeresésére. A 3. fejezetben segédfeladatként bevezetünk egy általános halmazlefedési feladatot, amely a metszősík módszer eddig megtett lépései során szerzett információkból épül fel. A segédfeladat kezelésével és megoldásával járó többletmunka árán lehetőség adódik a 2. fejezetben bemutatottnál hatékonyabb eljárás használatára. A 4. fejezet az előzőekben előállított határoló felületek levágására szolgáló eljárásokat tartalmaz. Az 5. fejezet az eredeti csúcspontból történő konvex metszést ismerteti. A 6. fejezet az előzőekben bemutatott eljárásokból felépíthető és az (1.1) feladat megoldására szolgáló metszősík módszert vázolja.

A 7. fejezet néhány olyan speciális konkáv minimalizálási feladatot ismertet, ahol az $f(x)$ függvénynek a P_0 -on való alulról korlátossága nyilvánvaló vagy könnyen ellenőrizhető. A 8. fejezet a módszerek implementációja során szerzett numerikus tapasztalatokat tartalmazza.

Köszönetet mondok BERNAU HEINZ-nek a dolgozat írása közben tett hasznos megjegyzéseiért.

2. Eredeti csúcspontot nem tartalmazó határoló felületek keresése

Legyen $N = \{1, \dots, n\}$ és $Z \subset N$ esetén

$$P_Z = \{x \in P_0 \mid x_j = 0, \quad j \in Z\}.$$

P_Z -t a P_0 egy *határoló felületének* nevezzük. Nyilván P_0 minden csúcspontja egyben egy 0 dimenziós határoló felület is. A metszősík módszer során eddig nyert k számú feltétel legyen a következő alakban megadva:

$$Dx + Ex_s = d$$

$$x_s \geq 0,$$

ahol D $k \times n$ -es mátrix, E $k \times k$ -s egységmátrix, $d, x_s \in R^k$, $x'_s = (x_{n+1}, \dots, x_{n+k})$, $S = \{n+1, \dots, n+k\}$. Legyen

$$Q_k = \{x \mid Dx + Ex_s = d, \quad x_s \geq 0\}$$

az eddigi metszősík feltételeknek eleget tevő n -dimenziós vektorok halmaza. Ekkor az aktuális poliéder $P_k = P_0 \cap Q_k$ és nyilván $P_k \subset P_{k-1} \subset \dots \subset P_0$. Tetszőleges P poliéder esetén jelölje $V(P)$ a P csúcspontjainak a halmazát. Ha valamely $x \in P_k$ -ra $x \in V(P_0)$ is teljesül, akkor az x pontot *eredeti csúcspontnak* nevezzük. Nyilván, ha $x \in P_k$ eredeti csúcspont, akkor $x \in V(P_k)$.

2.1. TÉTEL: Az $x^0 \in V(P_k)$ csúcspont akkor és csak akkor eredeti csúcspont, ha $x_s^0 = d - Dx^0$ esetén az

$$(2.1) \quad \begin{bmatrix} x^0 \\ x_s^0 \end{bmatrix}$$

megengedett bázismegoldás az

$$Ax = b$$

$$(2.2) \quad Dx + Ex_s = d$$

$$x \geq 0, \quad x_s \geq 0$$

feltétel rendszer egy olyan B megengedett bázisához tartozik, hogy $S \subset I_B$, ahol I_B a bázisindexek halmaza.

Bizonyítás: Ha $x^0 \in V(P_k)$ eredeti csúcspont, akkor (2.1) nyilván megengedett bázismegoldása (2.2)-nek. Az $x^0 \in V(P_0)$ csúcspontához megfeleltethető az $Ax = b$ feltétel rendszernek egy B_0 bázisa. Legyen $I_B = I_{B_0} \cup S$ és legyen B a (2.2) I_B -beli oszlopaiból álló $(m+k) \times (m+k)$ -s reguláris mátrix, amely így bázisa (2.2)-nek. Mivel B -hez éppen (2.1) tartozik, ezért B megengedett bázisa (2.2)-nek.

Fordítva, tegyük fel, hogy \mathbf{B} a (2.2) egy megengedett bázisa és $S \subset I_B$. Legyen (2.1) a \mathbf{B} -hez tartozó megengedett megoldás, nyilván $\mathbf{x}^0 \in P_0$ és $\mathbf{x}^0 \in V(P_k)$. Legyen $I_{B_0} = N \cap I_B$ és \mathbf{B}_0 az \mathbf{A} mátrix I_{B_0} -beli oszlopaiból álló $m \times m$ -es mátrix. Mivel a \mathbf{B} és \mathbf{B}_0 determinánsainak abszolút értéke egyenlő és nem nulla, így \mathbf{B}_0 az $\mathbf{A}\mathbf{x} = \mathbf{b}$ rendszer bázisa, amelynek éppen az \mathbf{x}^0 megengedett megoldás felel meg, tehát $\mathbf{x}^0 \in V(P_0)$.

Jelöljük (2.2) valamely \mathbf{B} bázisa esetén I_R -rel a nem bázisbeli változók indexhalmazát, azaz $I_R = N \cup S \setminus I_B$. Tetszőleges \mathbf{B} esetén I_R $n-m$ indexet tartalmaz, így egy eredeti csúcspont keresése (2.2) egy olyan \mathbf{B} megengedett bázisának keresését jelenti, amelyre $S \subset I_B$, azaz $N \cap I_R = n-m$ teljesül. A cél tehát (2.2) egy olyan megengedett bázisának a felkutatása, ahol maximális számú, $n-m$ N -beli változó nem bázisbeli. Induljunk ki (2.2) valamely \mathbf{B} megengedett bázisából és szimplex lépésekkel lépegezzünk (2.2) megengedett bázisain úgy, hogy közben a nem bázisbeli N -hez tartozó változók száma ne csökkenjen, sőt, lehetőleg növekedjen. Ez megtehető a következő, [9]-ből származó ötlettel. Válasszunk ki egy x_r , $r \in N \cap I_B$ változót, és az adott \mathbf{B} megengedett bázisból kiindulva oldjuk meg a

$$(2.3) \quad \min \{x_r | \mathbf{x} \in P_k\}$$

feladatot a szimplex módszernek azzal a módosításával, hogy csak S -hez tartozó változók léphetnek be a bázisba. Ha az optimális megoldásban az x_r változó 0 szinten bent van a bázisban, próbáljuk meg kipivotálni valamely S -hez tartozó nem bázisbeli változó felcserélésével.

A fenti belépési szabály úgy is felfogható, hogy az N -hez tartozó nem bázisbeli változókhoz tartozó oszlopokat ideiglenesen töröljük és ekkor a szokásos szimplex módszerrel dolgozhatunk. Ez egyben azt jelenti, hogy P_0 -nak az ezen változók által meghatározott határoló felületéről nem lépünk ki a (2.3) megoldása során. A belépési szabály biztosítja, hogy a megoldás közben a nem bázisbeli N -hez tartozó változók száma nem csökken, sőt, esetleg növekszik is.

A P_k poliédernek megfelelő (2.2) feltétel rendszer valamely \mathbf{B} bázisához tartozó kanonikus alak legyen a következőképpen megadva:

$$(2.4) \quad x_i + \sum_{j \in I_R} d_{ij} x_j = d_{i0}, \quad i \in I_B.$$

Ha a (2.3) feladat optimumértéke pozitív, akkor az optimális megengedett bázismegoldásban x_r bázisváltozó és az optimális \mathbf{B} bázishoz tartozó kanonikus alak x_r -hez rendelt sorában $d_{r0} > 0$ és $d_{rj} \leq 0$, $j \in I_R \cap S$. Ebből következik, hogy $Z = N \cap I_R$ esetén $P_Z \cap Q_k$ -n az x_r pozitív. Ha $d_{rj} \leq 0$, $j \in I_R \cap N$ is teljesül, akkor az x_r változó P_k -n csak pozitív értéket vesz fel.

Ha a (2.3) feladat optimumértéke nulla és az x_r változót nem sikerült kipivotálni valamely S -beli nem bázisváltozó felcserélésével, akkor az optimális kanonikus alakban $d_{r0} = 0$ és $d_{rj} = 0$, $j \in I_R \cap S$, továbbá ezek az együtthatók nullák is maradnak a (2.2) tetszőleges olyan megengedett bázisához tartozó kanonikus alakot tekintve, amelyet \mathbf{B} bázisból kiindulva, a speciális belépési szabályt használva szimplex lépésekkel elérünk.

Ha a (2.3) feladat megoldása során egy olyan megengedett bázismegoldáshoz jutunk, ahol minden x_j , $j \in S$ változó bázisváltozó, akkor megállhatunk, eredeti csúcspontot találtunk. Ha az optimális \mathbf{B} bázis nem eredeti csúcsponthoz tartozik, akkor kiválasztunk egy új x_r , $r \in N \cap I_B$ változót és megoldhatjuk az új (2.3) fel-

adatot a módosított belépési szabállyal, a \bar{B} bázisból kiindulva. Nyilvánvaló, hogy olyan változót, amelyet már tekintettünk, nem érdemes újra megvizsgálni, hiszen azon N -beli változók, amelyekről előzőleg megállapítottuk, hogy a módosított belépési szabály mellett a bázisban maradnak, továbbra sem lépnek ki a bázisból. A (2.3) feladat megoldásának egy adott lépésében legyen N_0 , illetve N_1 azon N -beli változók indexhalmaza, amelyekről kiderült, hogy nulla, illetve pozitív szinten a bázisban maradnak. Jelölje N_2 azon N -beli változók indexhalmazát, amelyekről tudjuk, hogy P_k -ban csak pozitív értéket vesznek fel. Nyilván $N_2 \subset N_1$. Legyen n_i az N_i indexhalmaz elemszáma, $i=0, 1, 2$.

2.1. LEMMA: Ha $n_2 > m$, akkor P_k nem tartalmaz eredeti csúcspontot.

Bizonyítás. Ha $j \in N_2$, akkor (2.2) tetszőleges B megengedett bázisa esetén $j \in I_B$, így az összes S -beli változó nem lehet bázisváltozó. A 2.1. tételből következik az állítás.

2.2. LEMMA. Tekintsük a (2.2) egy olyan B megengedett bázisát, amelyet a (2.3) feladat megoldása során értünk el. Legyen $Z = N \cap I_R$. Ekkor (2.2) tetszőleges megengedett bázismegoldása esetén a $Z \cup N_0$ indexhalmazba tartozó változók közül legalább n_0 számú bázisbeli, sőt, ha a megengedett bázismegoldás x része nincs rajta P_z -n, akkor legalább $n_0 + 1$ számú bázisbeli.

Bizonyítás. Tekintsük a B bázishoz tartozó szimplex táblát. Mivel $d_{r,j} = 0, r \in N_0, j \in (I_R \cap S) \cup \{0\}$, ezért a szimplex tábla

$$I_0 = \{0\} \cup (I_B \setminus N_0) \cup (I_R \cap S)$$

indexhalmazba tartozó oszlopaiból álló mátrix rangja $m + k - n_0$. Ezen mátrix rangja szimplex lépések végrehajtása során változatlan marad. A B bázisból kiindulva véges számú szimplex lépéssel eljuthatunk (2.2) bármely megengedett bázisához. Következésképpen (2.2) tetszőleges megengedett bázismegoldását tekintve az $I_0 \setminus \{0\}$ indexhalmazba tartozó változók közül legfeljebb $m + k - n_0$ lehet bázisbeli, de mivel

$$Z \cup N_0 \cup I_0 = N \cup S \cup \{0\},$$

ezért a $Z \cup N_0$ indexhalmazba tartozó változók közül legalább n_0 számú bázisbeli. Ha a megengedett bázismegoldás x része nincs rajta P_z -n, akkor legalább egy Z -hez tartozó változó pozitív szinten bent van a bázisban. Ekkor azonban az $I_0 \setminus \{0\}$ indexhalmazba tartozó változók közül nem lehet $m + k - n_0$ számú bázisváltozó, mivel a megfelelő szimplex tábla I_0 -hoz tartozó oszlopaiból álló mátrix rangja $m + k - n_0 + 1$ lenne. Innen következik állításunk második része is.

2.3. LEMMA. Tekintsük a (2.2) egy olyan B megengedett bázisát, amelyet a (2.3) feladat megoldása során értünk el. Legyen $Z = N \cap I_R$ és $\hat{l}_1 = n_0 + n_1 - m$. Ha $\hat{l}_1 > 0$, akkor $P_Z \cap Q_k$ -ban nincs eredeti csúcspont. Ha $\hat{l}_1 \geq 0$, akkor tetszőleges olyan P_k -beli eredeti csúcspont, amely nincs rajta P_z -n, hozzárendelhető a (2.2) egy olyan megengedett bázisa, amelynél az N_1 -hez tartozó változók közül legalább $\hat{l}_1 + 1$ számú nincs bent a bázisban.

Bizonyítás. A (2.2) tetszőleges olyan megengedett bázismegoldása esetén, amelynek x része a P_z -n fekszik, a $Z \cup N_0 \cup N_1$ indexhalmazba tartozó változók közül

legalább $n_0 + n_1$ számú bázisbeli. Így $\hat{f}_1 > 0$ esetén a 2.1. tételből következik az első állítás. Ha egy P_k -ban levő eredeti csúcsponthoz rendelt megengedett bázismegoldás x része nincs P_Z -ben, akkor a $Z \cup N_0$ -beli változók közül legalább $n_0 + 1$ bázisváltozó. Mivel ehhez a megengedett bázismegoldáshoz (2.2) olyan bázisa tartozik, hogy pontosan m számú N -beli változó bázisváltozó, ezért következik a második állítás is.

A 2.3. lemma alapján következik, hogy a (2.3) feladat megoldása közben leállhatunk, mielőtt $m+1$ számú változóról kiderült, hogy a módosított belépési szabály mellett bázisban marad. Ez azt jelenti, hogy a nem bázisbeli, N -hez tartozó változók által meghatározott P_Z határoló felület és az eddigi metszéseknek eleget tevő pontok Q_k halmazának közös részén nincs már eredeti csúcspont. Ha van még eredeti csúcspont P_k -ban, akkor ahhoz (2.2) olyan megengedett bázismegoldása tartozik, hogy az N_1 -be tartozó változók közül legalább kettő bázison kívüli, így nulla szintű. Az N_2 -beli változók P_k -n pozitívak, tehát $n_1 - n_2 \leq 1$ esetén P_k nem tartalmaz eredeti csúcspontot. Az eddigi eredményeket a következő algoritmusban foglalhatjuk össze, amely egy eredeti csúcspontot vagy egy eredeti csúcspontot már nem tartalmazó határoló felületet talál.

2.1. Algoritmus

1. Legyen $N_0 = \emptyset$, N_2 pedig olyan N -beli változók indexhalmaza, amelyekről már tudjuk, hogy P_k -n pozitívak. Legyen $N_1 = N_2$. Tekintsük (2.2) tetszőleges B megengedett bázisát.
2. Ha $S \subset I_B$, vagy $n_0 + n_1 > m$, akkor menjünk 5.-re, különben válasszunk ki egy $r \in (N \cap I_B) \setminus (N_0 \cup N_1)$ indexet.
3. Az utoljára elért megengedett bázisból kiindulva oldjuk meg a

$$\min \{x_r | x \in P_k\}$$

feladatot a szimplex módszernek azzal a módosításával, hogy csak S -beli változók léphetnek be a bázisba.

4. Ha az optimális megengedett bázismegoldásban $x_r > 0$, akkor r -t csatoljuk az N_1 indexhalmazhoz, sőt, ha az optimális szimplex táblában $d_{rj} \leq 0$, $j \in N \cap I_R$ is teljesül, akkor az N_2 indexhalmazhoz is. Ha $x_r = 0$ és x_r bázisváltozó, akkor próbáljuk meg kipivotálni valamely $I_R \cap S$ -beli változóval felcserélve. Ha ez nem sikerül, akkor r -t csatoljuk az N_0 indexhalmazhoz. Menjünk 2.-re,
5. Legyen $Z = N \cap I_R$. Ha $S \subset I_B$, akkor P_Z egy eredeti csúcspont, különben $P_Z \cap Q_k \cap V(P_0) = \emptyset$ és tetszőleges P_k -ban levő eredeti csúcsponthoz olyan megengedett bázismegoldás rendelhető, hogy legalább kettő $N_1 \setminus N_2$ indexhalmazba tartozó változó bázison kívül van, tehát $n_1 - n_2 \leq 1$ esetén $P_k \cap V(P_0) = \emptyset$.

Legyen

$$Z(x) = \{j | x_j = 0\},$$

ahol $x \in R^n$. Tegyük fel, hogy valamely P_Z határoló felület esetén $P_Z \cap Q_k \neq \emptyset$. Azt mondjuk, hogy a P_Z határoló felület a P_0 poliédernek egy *extremális határoló*

felülete a Q_k halmazra vonatkozóan, ha bármely két $x^1, x^2 \in P_Z \cap Q_k$ pont esetén $Z(x^1) = Z(x^2)$ [9]. Ez azt jelenti, hogy $P_Z \cap Q_k$ nem tartalmaz pontot P_0 -nak alacsonyabb dimenziós határoló felületéből. Nyilvánvaló, hogy minden P_k -ba eső eredeti csúcspont egyben extrémális határoló felület is Q_k -ra vonatkozóan. Ha egy extrémális határoló felület 0 dimenziós, akkor eredeti csúcspont, egyébként nem tartalmaz eredeti csúcspontot.

Az extrémális határoló felületet kereső *Majthay—Whinston-módszer* [9] a 2.1. algoritmushoz hasonlóan működik. Addig old meg (2.3) feladatokat a módosított belépési szabállyal, amíg az összes, bázisban levő, N -hez tartozó változót meg nem vizsgálta. Legyen B az ilyen módon utoljára elért megengedett bázis. Ekkor a következő állítható [9]. Ha $N \subset I_B$, akkor P_0 -nak nincs extrémális határoló felülete a Q_k -ra vonatkozóan, egyébként $Z = I_R \cap N$ esetén P_Z egy ilyen határoló felület. Ha az utóbbi esetben $I_R \subset N$ is teljesül, akkor P_Z egy eredeti csúcspont. Tekintsük a B által meghatározott megengedett megoldásban a pozitív N -beli komponensek indexhalmazát. Ekkor tetszőleges a P_k -ban levő, de nem a P_Z -n fekvő eredeti csúcspontot tekintve legalább egy, ebbe az indexhalmazba tartozó komponens nulla.

Könnyen látható, hogy az eredeti csúcspont keresése szempontjából a 2.1. algoritmus hatékonyabb. Ha egy P_Z határoló felület esetén $P_Z \cap Q_k \cap V(P_0) = \emptyset$, attól még P_Z vagy részfelületei extrémális határoló felületek lehetnek Q_k -ra vonatkozóan. Egy ilyen határoló felületen (1.1) célfüggvénye nem vesz fel jobb értéket az eddig talált legjobb célfüggvényértéknél, így az ezen felületen történő további lépegetés felesleges.

Mind a 2.1 algoritmusnál, mind a [9]-ben közölt módszernél felmerül a kérdés, hogy a következő (2.3) feladat előállításához melyik x_j változót válasszuk ki a lehetségesek közül. [9] erre nem ad tanácsot. [15, 17] a legnagyobb értékű x_j kijelölését javasolja. Az implementációk során általunk szerzett numerikus tapasztalatok a minimális értékű x_j választását sugallják. Ezt azzal a heurisztikus érveléssel is alátámaszthatjuk, hogy a (2.3) feladatot valószínűleg kevesebb lépéssel oldjuk meg, ha kisebb célfüggvényértékű megengedett bázisról indulunk.

Tudjuk, hogy a 2.1. algoritmus során az $N_0 \cup N_1$ indexhalmazhoz tartozó változók a bázisban maradnak. Ekkor egy szimplex lépés végrehajtásakor a bázisba belépő x_k változó ismeretében a bázisból kilépő x_l változó a következőképpen nyerhető:

$$(2.5) \quad \min \{d_{i0}/d_{ik} | d_{ik} > 0, \quad i \in I_B \setminus (N_0 \cup N_1)\} = d_{i0}/d_{ik}.$$

3. A bázisban maradó N -beli változók számának egy alsó becslése

Tegyük fel, hogy a P_k -beli pontoknak ki kell elégíteniük a

$$(3.1) \quad \sum_{j \in N} \alpha_j x_j = \alpha_0$$

feltételt, ahol $\alpha_0 > 0$. Legyen

$$J^+ = \{j \in N | \alpha_j > 0\}.$$

Ekkor legalább egy J^+ -hoz tartozó x_j változónak pozitívnak, így (2.2) bármely megengedett bázismegoldása esetén bázisváltozónak kell lennie. Vezessük be az

$y = (y_1, \dots, y_n)$ 0—1 változókból álló vektort, ahol

$$y_j = \begin{cases} 1, & \text{ha } x_j \text{ bázisváltozó} \\ 0, & \text{ha } x_j \text{ nem bázisváltozó} \end{cases} \quad (j = 1, \dots, n)$$

Ekkor (2.2) tetszőleges megengedett bázismegoldását tekintve fennáll a

$$(3.2) \quad \sum_{j \in J^+} y_j \equiv 1$$

egyenlőtlenség. Sőt, mivel az újabb l számú feltétel csatolásával nyert P_{k+l} poliéder része P_k -nak, így a P_{k+l} -nek megfelelő (2.2) feltételrendszer bármely megengedett bázismegoldása is teljesíti (3.2)-t. Ha (3.1)-ben = helyett \equiv áll, akkor (3.2) ugyanúgy teljesül. Ha (3.1)-ben = vagy \equiv jel van és $\alpha_0 < 0$, akkor a

$$J^- = \{j \in N \mid \alpha_j < 0\}$$

jelöléssel

$$(3.3) \quad \sum_{j \in J^-} y_j \equiv 1$$

egyenlőtlenség áll fenn a P_{k+l} ($l=0, 1, \dots$) csúcspontjaihoz tartozó megengedett bázismegoldásokra.

Ha a 2.1. algoritmus során valamely x_r , $r \in I_B \cap N$ változóról kiderül, hogy a $Z = N \cap I_R$ által meghatározott P_Z határoló felület Q_k -ba eső részén pozitív, azaz $r \in N_1$, akkor a megfelelő (2.3) feladat optimális kanonikus alakjának x_r -hez tartozó sorában $d_{rj} \equiv 0$, $j \in S \cap I_R$ és $d_{r0} > 0$, így fennáll a

$$(3.4) \quad x_r + \sum_{j \in I_R \cap N} d_{rj} x_j \equiv d_{r0}$$

összefüggés, amiből elkészíthető a megfelelő (3.2) egyenlőtlenség is. Ha $d_{rj} \equiv 0$, $j \in N \cap I_R$ is teljesül, akkor nyilván $y_r \equiv 1$, azaz $y_r = 1$.

Tegyük fel, hogy a 2.1. algoritmus során valamely x_r , $r \in N \cap I_B$ változóról kiderül, hogy az adott belépési szabály mellett nulla szinten bennmarad a bázisban, azaz $r \in N_0$. Ekkor a 2.2. lemma szerint a P_k poliéder csúcspontjaihoz tartozó bázismegoldásokra érvényes a

$$(3.5) \quad \sum_{j \in Z \cup N_0} y_j \equiv n_0$$

összefüggés, sőt, olyan csúcspont esetén, amely nem a P_Z határoló felületen fekszik a

$$(3.6) \quad \sum_{j \in Z \cup N_0} y_j \equiv n_0 + 1$$

egyenlőtlenség is.

Tekintsünk egy $x^0 \in P_k$ eredeti csúcspontot. Ekkor x^0 csúcspontja minden P_l , $0 \leq l \leq k-1$ poliédernek is. Tudjuk, hogy az eredeti csúcsponthoz megfeleltethető a (2.2) feltételrendszernek egy olyan bázismegoldása, ahol az S -beli változók bázisban vannak, így a P_k -nak megfelelő (2.2) rendszer x^0 -hoz tartozó bázisából előállítható a P_l -nek ($0 \leq l \leq k-1$) megfelelő (2.2) rendszer x^0 -hoz tartozó bázisa. Tehát, ha a P_l poliéder csúcspontjaihoz tartozó bázismegoldásokra fennáll egy

$$(3.7) \quad \sum_{j \in U_i} y_j \equiv \beta_i$$

összefüggés, ahol $U_i \subset N$, akkor $l \leq k$ esetén a P_k -beli eredeti csúcspontokhoz tartozó bázismegoldásokra is teljesül (3.7). Gyűjtsük össze a P_k -ban levő eredeti csúcspontokra fennálló (3.7) feltételeket:

$$\sum_{j \in U_i} y_j \cong \beta_i, \quad i = 1, \dots, p$$

$$y_j \in \{0, 1\}, \quad j = 1, \dots, n$$

ahol $U_i \subset N$ és β_i -ről feltehető, hogy pozitív egész szám.

3.1. LEMMA. Ha P_k -ban van eredeti csúcspont, akkor a következő általános halmazlefedési feladat optimumértéke nem nagyobb m -nél:

$$\min \sum_{j=1}^n y_j$$

$$(3.8) \quad \sum_{j \in U_i} y_j \cong \beta_i, \quad i = 1, \dots, p$$

$$y_j \in \{0, 1\} \quad j = 1, \dots, n.$$

Bizonyítás. A 2.1. tételből következik, hogy P_k -beli eredeti csúcspontokhoz a megfelelő (2.2) rendszer olyan megengedett bázismegoldása is tartozik, hogy az $x_j, j \in N$ változók közül m számú bázisváltozó. Az előálló y vektor nyilván kielégíti (3.8) feltételeit és

$$\sum_{j=1}^n y_j = m.$$

Ebből következik az állítás.

A metszősík módszer során új feltételeket csatolunk (3.8)-hoz, ezért az új feladatok állandó egzakt megoldása nagyon időigényes lenne. Több gyors, heurisztikus alsó, illetve felső becslés ismert azonban (3.8) optimumértékére [7, 14]. Legyen \hat{u} egy ilyen felső becslés. Ha $\hat{u} \leq m$, akkor nem állíthatjuk a 3.1. lemma felhasználásával, hogy P_k -ban nincs eredeti csúcspont. Ha $\hat{u} > m$, akkor tekintsünk egy \hat{l} alsó becslést (3.8) optimumértékére vonatkozóan. Ha $\hat{l} > m$, akkor $P_k \cap V(P^0) = \emptyset$, azaz P_k -ban nincs eredeti csúcspont. Egy új feltétel csatolásakor az új, illetve a régi (3.8) feladat optimumértékei legfeljebb az új feltétel jobb oldalán álló mennyiséggel térhetnek el egymástól. Ha valamely (3.8) feladatra $\hat{u} < m$ adódik, akkor mindaddig, amíg az \hat{u} és az ezután csatolandó új feltételek jobb oldalának összege nem haladja meg az m értéket, az alsó, illetve a felső becslések kiszámítása felesleges.

3.2. LEMMA: Tekintsük a 2.1. algoritmus valamely lépéséhez tartozó megengedett bázist. Legyenek a Z, N_0, N_1 indexhalmazok az ottaniak szerint meghatározva. Legyen \hat{l}_1 a következő általános halmazlefedési feladat optimumértékének egy alsó becslése:

$$\min \sum_{j=1}^n y_j$$

$$(3.9) \quad \sum_{j \in U_i} y_j \cong \beta_i, \quad i = 1, \dots, p$$

$$y_j = 1, \quad j \in N_1$$

$$y_j \in \{0, 1\} \quad j \in N \setminus N_1.$$

Ha $\hat{l}_1 > m$, akkor $P_Z \cap Q_k \cap V(P^0) = \emptyset$. Ebben az esetben legyen \hat{l}_2 a

$$(3.10) \quad \sum_{j \in Z \cup N_0} y_j \cong n_0 + 1$$

feltételnek (3.9)-hez történő csatolásával nyert általános halmazlefedési feladat optimumértékének egy alsó becslése. Nyilván feltehető, hogy $\hat{l}_2 \cong \hat{l}_1$. Ekkor tetszőleges P_k -beli eredeti csúcsponthoz megfeleltethető (2.2) olyan megengedett bázis-megoldása, hogy az $N_1 \setminus N_2$ indexhalmazhoz tartozó változók közül legalább $\hat{l}_2 - m$ számú nem bázisváltozó.

Bizonyítás. A (3.9) feladat feltételrendszere a P_k -beli eredeti csúcspontokra fennálló (3.8) feltételek mellett a P_Z határoló felületen levő csúcspontokra nyilvánvalóan teljesülő $y_j = 1, j \in N_1$ feltételeket tartalmazza. Ha $\hat{l}_1 > m$, akkor (3.9) optimumértéke is nagyobb m -nél, így a 3.1. lemmához hasonlóan bizonyítható az első állítás. Ebben az esetben, ha P_k -ban van eredeti csúcspont, akkor az nincs rajta P_Z -n, így a 2.2. lemma alapján erre a csúcspontra teljesül (3.10). A (3.9)–(3.19) feladat optimumértéke nem kisebb a (3.9) feladaténál, ezért $\hat{l}_2 \cong \hat{l}_1 > m$ feltehető. Ahhoz, hogy az eredeti csúcsponthoz rendelhető y vektorra

$$\sum_{j=1}^n y_j = m$$

teljesüljön, szükséges, hogy legalább $\hat{l}_2 - m$ számú $j \in N_1$ esetén ne álljon fenn $y_j = 1$, azaz az $N_1 \setminus N_2$ indexhalmazhoz tartozó x_j változók közül legalább $\hat{l}_2 - m$ számú ne legyen bázisváltozó.

A metszősík módszer során újabb és újabb (3.7) feltételeket állítunk elő, ezért a (3.8) és (3.9) feladatok feltételeinek száma növekszik. Ismeretes azonban több olyan méretcsökkentő eljárás, amelynek általános halmazlefedési feladatra történő alkalmazásával egy kisebb méretű, de a kiindulási feladat optimumértékével azonos optimumértékű általános halmazlefedési feladatot nyerhetünk. Ezek ismertetésére itt nem térünk ki, részletes tárgyalása található pl. [2, 7, 14]-ben.

Megjegyezzük, hogy a 3.1. lemma a 2.1. lemma általánosítása, mivel $n_2 > m$ esetén a (3.8) feladat optimumértéke nyilván nagyobb m -nél. A 3.2. lemma a 2.3. lemma általánosítása, mivel $n_0 + n_1$ a (3.9), $n_0 + n_1 + 1$ pedig a (3.9)–(3.10) feladat optimumértékének nyilvánvaló alsó becslése. Ilyen módon az itt bemutatandó 3.1. algoritmus a 2.1. algoritmus általánosításaként tekinthető, mivel a segédfeladat kezelésével és optimumértékének becslésével járó többletmunka árán a 2.1. algoritmushoz szükségesnél esetleg lényegesen kevesebb szimplex lépéssel dönthet arról, hogy egy adott határoló felületen van-e eredeti csúcspont.

3.1. Algoritmus

1. Azonos a 2.1. algoritmus 1. lépésével, azzal a kiegészítéssel, hogy legyen a (3.8) feladat feltételrendszere a P_k -beli eredeti csúcspontokra fennálló (3.7) feltételek halmaza.
2. Ha $S \subset I_B$, akkor menjünk 5.-re. Legyen \hat{l} a (3.8) feladat optimumértékének alsó becslése. Ha $\hat{l} > m$, akkor $P_k \cap V(P_0) = \emptyset$, tehát leállhatunk, amúgy

legyen \hat{l}_1 a (3.9) feladat optimumértékének alsó becslése. Ha $\hat{l}_1 > m$, akkor legyen \hat{l}_2 a (3.9)–(3.10) feladat optimumértékének alsó becslése és menjünk 5.-re, különben válasszunk ki egy $r \in (N \cap I_B) \setminus (N_0 \cup N_1)$ indexet.

3. Azonos a 2.1. algoritmus 3. lépésével.
4. Azonos a 2.1. algoritmus 4. lépésével a következő kiegészítésekkel. Az $r \in N_1$ esetben a (3.4) egyenlőtlenségből előálló (3.2) feltételt, az $r \in N_0$ esetben pedig a (3.5) feltételt csatoljuk (3.8)-hoz.
5. Legyen $Z = N \cap I_R$. Ha $S \subset I_B$, akkor P_Z egy eredeti csúcspont, különben $P_Z \cap Q_k \cap V(P_0) = \emptyset$ és tetszőleges P_k -ban levő eredeti csúcspont-hoz olyan megengedett bázismegoldás rendelhető, hogy legalább $\hat{l}_2 - m$ számú $N_1 \setminus N_2$ indexhalmazba tartozó változó bázison kívül van, tehát $n_1 - n_2 \leq \hat{l}_2 - m$ esetén $P_k \cap V(P_0) = \emptyset$.

A 3.1. algoritmus során egy feltétel (3.8)-hoz történő csatolásakor vagy a (3.9) feladatnak (3.8)-ból való előállításakor érdemes végrehajtani a méretcsökkentő eljárásokat. Ugyancsak célszerű a megfelelő általános halmazlefedési feladatok optimumértéke felső becslésének fentebb tárgyalt használata.

4. Eredeti csúcspontot már nem tartalmazó határoló felületek levágása

Tegyük fel, hogy a 2.1. vagy 3.1. algoritmus felhasználásával valamely P_Z határoló felületről kiderítettük, hogy Q_k -ba eső részén nincs eredeti csúcspont. Ennek a határoló felületnek a pontjait tehát nem érdemes vizsgálnunk az (1.1) feladat megoldását célzó eljárás további lépéseiben. Ezt a következőképpen érhetjük el. Konstruálunk egy olyan hipersíkot, hogy a P_Z pontjai a hipersík által meghatározott egyik nyílt féltérbe, az esetleg P_k -ban levő eredeti csúcspontok pedig a másik zárt féltérbe esnek. Az utóbbi zárt féltérnek feltétel alakjában (2.2) eddigi feltételeihez való csatolásával P_Z pontjait kizárjuk a további keresésből. A hipersík előállítására azt az információt használjuk fel, hogy a $P_Z \cap Q_k$ -n csak pozitív értéket felvevő változók közül bizonyos számúnak bázison kívül, így nulla szinten kell lennie P_k -beli eredeti csúcspont-hoz tartozó bázismegoldás esetén. Jelöljük J -vel ezen változók indexhalmazát és tegyük fel, hogy l számú változónak kell nullává válnia a J indexhalmazból. A 2.1., illetve a 3.1. algoritmus használata esetén $J = N_1 \setminus N_2$. Az előbbi esetben $l = 2$, az utóbbi esetben $l = \hat{l}_2 - m$. Nyilván feltehetjük, hogy $l \leq |J|$. Az itt bemutatandó két eljárás az extrémális határoló felületet levágó parametrikus [9], illetve diszjunktív [15] metszés módosítása és kiegészítése a 2.1. és a 3.1. algoritmus által előállított határoló felületek levágására.

4.1. Parametrikus metszés

Válasszunk tetszőleges $p'_j > 0, j \in Z$ számokat és legyen

$$H_Z(p') = \{x | x \in R^m, \sum_{j \in Z} p'_j x_j \leq p'\}.$$

Tekintsük a következő, p' paramértől függő lineáris programozási feladatot:

$$(4.1) \quad \min \{x_r | x \in P_k \cap H_Z(p')\},$$

ahol $r \in J$. A feladat optimumértéket jelöljük $\pi(r, p')$ -vel. Mivel $P_k \cap H_Z(0) = P_Z \cap Q_k$, így $\pi(r, 0) > 0$ tetszőleges $r \in J$ esetén. $\pi(r, p')$ a p' szakaszonként lineáris konvex függvénye a nemnegatív p' számok halmazán. Legyen

$$p_r = \min \{p' | \pi(r, p') = 0, p' \geq 0\}.$$

A p_r meghatározása egy parametrikus lineáris programozási feladat megoldását jelenti. Nyilvánvaló, hogy p_r akkor és csak akkor $+\infty$, ha x_r a p_k poliéderen pozitív. Ha valamely x_r változóra $p_r = +\infty$ adódik, akkor az r indexet csatolhatjuk az N_2 indexhalmazhoz. Legyen \bar{p} a $p_r, r \in J$ számok közül az l -edik legkisebb. Ha $\bar{p} = +\infty$, akkor $P_k \cap V(P_0) = \emptyset$, egyébként állítsuk elő a

$$(4.2) \quad \sum_{j \in Z} p'_j x_j \cong \bar{p}$$

feltételt. Tetszőleges P_k -ban levő eredeti csúcspont teljesíti (4.2)-t, viszont $\bar{p} > 0$ miatt a P_Z határoló felület pontjaira nem áll fenn (4.2).

4.2. Diszjunktív metszés

Mindegyik J -beli x_r változóról egy (2.3) feladat megoldásával derült ki, hogy $P_Z \cap Q_k$ -n pozitív, ezért tekintsük a megfelelő (2.3) feladat optimális bázisából előálló kanonikus alak x_r -hez tartozó sorát:

$$x_r + \sum_{j \in I_R} d_{rj} x_j = d_{r0}.$$

Nyilván $d_{rj} \leq 0, j \in S \cap I_R$ és $d_{r0} > 0$. A (2.2) tetszőleges olyan megengedett megoldására, amelynél x_r eltűnik, fennáll a következő egyenlőtlenség:

$$\sum_{j \in I_R} (d_{rj}/d_{r0}) x_j \cong 1.$$

Az előző egyenlőtlenséget megfelelő 0 együtthatók bevezetésével írjuk a következő alakra:

$$\sum_{j \in N \cup S} \alpha_{rj} x_j \cong 1.$$

Nyilván $j \notin Z$ esetén $\alpha_{rj} \leq 0$, de mivel J -be nem vettünk bele olyan indexet, amelyre $d_{rj} \leq 0, j \in I_R$ teljesülne, így van legalább egy $\alpha_{rj} > 0, j \in Z$ együttható. A (2.2) feltételrendszer tetszőleges eredeti csúcsponthoz tartozó megoldása legalább l számú $r \in J$ esetén teljesíti a következő feltételeket:

$$(4.3) \quad \left\{ \sum_{j \in N \cup S} \alpha_{rj} x_j \cong 1, x_j \geq 0, j \in N \cup S \right\}.$$

Hagyjunk el a J indexhalmazból tetszőlegesen $l-1$ elemet. Jelöljük az így nyert részhalmazt J^* -gal. A (2.2) bármely eredeti csúcsponthoz tartozó megoldása legalább egy $r \in J^*$ esetén kielégíti a megfelelő (4.3) feltételrendszert. Legyen

$$\alpha_j^* = \max_{r \in J^*} \alpha_{rj}, \quad j \in N \cup S.$$

A P_k -beli eredeti csúcspontok nyilván eleget tesznek a

$$(4.4) \quad \sum_{j \in N \cup S} \alpha_j^* x_j \cong 1$$

egyenlőtlenségeknek, azonban $\alpha_j^* \leq 0$, $j \notin Z$ miatt (4.4) nem teljesül P_Z pontjaira. A (4.4) metszés mélyebbé tehető [5]. Legyen

$$\gamma_r = \min \{ \alpha_{rj}^* / \alpha_{rj} \mid \alpha_{rj} > 0, j \in N \cup S \}, \quad r \in J^*$$

és

$$\alpha_j^{**} = \max_{r \in J^*} \gamma_r \alpha_{rj}, \quad j \in N \cup S.$$

Könnyen látható [17], hogy

$$\alpha_j^{**} \begin{cases} = \alpha_j^*, & \text{ha } \alpha_j^* \cong 0 \\ \cong \alpha_j^*, & \text{ha } \alpha_j^* < 0, \end{cases}$$

és a

$$(4.5) \quad \sum_{j \in N \cup S} \alpha_j^{**} x_j \cong 1$$

feltétel is elválasztja a P_k -beli eredeti csúcspontokat a P_Z felülettől. Minél kisebbek az α_j^{**} együtthatók, annál mélyebb a (4.5) metszés. Ezért J^* -nak J -ből való előállításakor olyan $r \in J$ indexeket érdemes elhagyni, amelyekhez tartozó sorokban található a legnagyobb α_{rj} együtthatók. Megkeressük a maximális α_{rj} , $r \in J$, $j \in N \cup S$ együtthatót, töröljük a megfelelő r -t J -ből, és így tovább összesen $(l-1)$ -szer.

Megjegyezzük, hogy $|J|=l$ esetén a (4.5) metszés előállítása nem szükséges. Tudjuk, hogy ekkor bármely $x \in P_k$ eredeti csúcsponthoz megfeleltethető a (2.2) olyan bázisa, hogy a J -beli változók nincsenek benne a bázisban, így

$$(4.6) \quad \sum_{j \in J} x_j = 0,$$

azaz $x \in P_J$. Mivel $x \in P_Z \cap Q_k$ esetén $x_j > 0$, $j \in J$, ezért a (4.6) egyenlőség csatolása kizárja P_Z pontjait a további keresésből. Azonban a (4.6) feltétel csatolásától eltekinthetünk, hiszen a P_k -beli eredeti csúcspontokat elég (2.2) olyan megengedett bázismegoldásai között keresni, amelyek teljesítik a

$$(4.7) \quad \text{„}J\text{-beli változó nincsen bent a bázisban”}$$

feltételt. A (2.2) rendszer (4.7)-nek is eleget tevő megengedett bázisának előállítása a kétfázisú simplex módszer I. fázisával azonos módon történik. Ha ilyen bázist nem tudunk elérni, akkor nincs eredeti csúcspont P_k -ban. Ha eljutunk egy megfelelő bázishoz, akkor véglegesen törölhetjük (2.2) J -hez tartozó oszlopait és így a (4.7) feltétel is elhagyható. A parametrikus metszés esetén is alkalmazhatjuk az előbbi ötletet. Legyen J^* azon $r \in J$ indexek halmaza, amelyekre p_r véges. Ekkor $|J^*|=l$ esetén, J helyett a J^* indexhalmazt használva, az előbbiek szerint járhatunk el.

A (4.2) vagy (4.5) egyenlőtlenségek (2.2)-höz csatolása előtt hozzuk azokat a (2.2) által megkívánt alakra. Ha (4.5)-ben valamely x_{n+j} változó ($1 \leq j \leq k$) $\alpha_{n+j}^{**} \neq 0$ együtthatóval szerepel, akkor (2.2) $(m+j)$ -edik feltétele α_{n+j}^{**} -szerezésének (4.5)-ből történő kivonásával nyert új feltételben az x_{n+j} együtthatója nulla. A megfelelő irányú egyenlőtlenséget (-1) -gyel való szorzással érhetjük el.

5. Konvex metszés előállítása eredeti csúcspontból

Tekintsük most azt az esetet, amikor a 2.1. vagy 3.1. algoritmus során egy P_k -ban levő eredeti csúcsponthoz jutunk. Egy metszősík feltételt állítunk elő úgy, hogy a levágandó részen a célfüggvény ne vegyen fel az eddig talált legjobbnál jobb értéket és a metszés az eredeti csúcspontot is kizárja. Egy ilyen metszés előállításakor figyelembe vesszük az $f(\mathbf{x})$ célfüggvényt, szemben az előző fejezetben bemutatott metszésekkel, ahol csak a P_k poliédert meghatározó lineáris feltételeket használtuk fel.

Legyen \mathbf{x}^0 a P_k tetszőleges nemdegenerált csúcspontja, \mathbf{B} a (2.2) \mathbf{x}^0 -hoz tartozó bázisa és tekintsük a \mathbf{B} -nek megfelelő kanonikus alakot. A (2.2) rendszer

$$\hat{\mathbf{x}}^0 = \begin{bmatrix} \mathbf{x}^0 \\ \mathbf{x}_s^0 \end{bmatrix}$$

extremális pontjából kiinduló élek $\hat{\mathbf{z}}^j, j \in I_R$ irányai a következőképpen írhatók fel:

$$\hat{\mathbf{z}}_k^j = \begin{cases} -d_{kj}, & \text{ha } k \in I_B \\ 1, & \text{ha } k = j \\ 0, & \text{különben.} \end{cases}$$

Legyen

$$I_R^{(1)} = \{j \in I_R \mid d_{ij} \leq 0, \quad i \in I_B\}$$

és

$$I_R^{(2)} = I_R \setminus I_R^{(1)}.$$

Ekkor az $\hat{\mathbf{x}}^0$ -ból kiinduló $\hat{\mathbf{z}}^j, j \in I_R^{(1)}$ irányú élek nem korlátosak, míg a $\hat{\mathbf{z}}^j, j \in I_R^{(2)}$ irányúak korlátosak és a

$$\theta_j = \frac{d_{i(j)0}}{d_{i(j)j}} = \min \left\{ \frac{d_{i0}}{d_{ij}} \mid d_{ij} > 0, \quad i \in I_B \right\}, \quad j \in I_R^{(2)}$$

által meghatározott

$$\hat{\mathbf{x}}^j = \hat{\mathbf{x}}^0 + \theta_j \hat{\mathbf{z}}^j, \quad j \in I_R^{(2)}$$

pontok az adott élek másik végpontjai. Jelölje f^0 a metszősík módszer során eddig talált legjobb célfüggvényértéket, amely meghatározásánál $f(\mathbf{x})$ -nek az $\hat{\mathbf{x}}^0$ és $\hat{\mathbf{x}}^j, j \in I_R^{(2)}$ pontokban felvett értékét is figyelembe vettük, a függvény argumentumául az első m számú komponens tekintve. Határozzuk meg a

$$\lambda_j = \sup \{ \lambda \mid f(\hat{\mathbf{x}}^0 + \lambda \hat{\mathbf{z}}^j) \leq f^0 \}, \quad j \in I_R$$

számokat. Ha $j \in I_R^{(1)}$, akkor $\lambda_j = +\infty$, mivel $f(\mathbf{x})$ alulról korlátos a P_0 poliéderen, így P_k -n is és $f(\mathbf{x})$ konkáv. Ha $j \in I_R^{(2)}$, akkor nyilván $\lambda_j \leq \theta_j$. Legyen

$$t_j = \begin{cases} 1/\lambda_j, & \text{ha } \lambda_j < +\infty \\ 0, & \text{ha } \lambda_j = +\infty \end{cases} \quad j \in I_R.$$

Ekkor

$$\mathbf{x} \in P_k \cap \{ \mathbf{x} \in R^n \mid \sum_{j \in I_R} t_j x_j \leq 1 \}$$

esetén $f(\mathbf{x}) \equiv f^0$ [2, 4, 18]. Ennek az összefüggésnek a bizonyításánál az $f(\mathbf{x})$ kvázikonkáv volta, ami konkáv függvény esetén nyilván teljesül, kerül csak felhasználásra. Csatoljuk a

$$(5.1) \quad \sum_{j \in I_R} t_j x_j \equiv 1$$

feltételt (2.2)-hez. Nyilvánvaló, hogy $\hat{\mathbf{x}}^0$ nem teljesíti (5.1)-t, így az \mathbf{x}^0 csúcsot levágtuk a P_k poliéderről. Az (5.1) metszést *konvex metszésnek* nevezik, de szokásos a *TUY-metszés* elnevezés is. A fentebbi megjegyzésből következik, hogy (5.1) használható kvázikonkáv függvény korlátos poliéderen történő minimalizálásakor. A poliéder korlátozottságát itt azért kötjük ki, mert nem korlátos poliéder és alulról korlátos kvázikonkáv függvény esetén az infimumérték nem vétetik fel feltétlenül, gondoljunk csak a nemnegatív valós számok halmazára és az e^{-x} függvényre.

Degenerált \mathbf{x}^0 csúcspont esetén is előállítható a fenti követelményeknek eleget tevő konvex metszés [2, 4]. Ebben az esetben is szükség van azonban a csúcspontból kiinduló élek irányának meghatározására, ami a degeneráció miatt nagyon munkaigényes lehet [10]. A degenerált csúcspontból történő konvex metszés részletes tárgyalása található [4, 9]-ben.

Egy \mathbf{x}^0 degenerált eredeti csúcspont levágásának másik módja származtatható abból a tényből, hogy \mathbf{x}^0 a P_0 poliéder egy extrémális határoló felülete Q_k -ra vonatkozóan. Legyen

$$J = \{j \in N \mid x_j^0 > 0\}.$$

Ekkor tetszőleges P_k -ban levő \mathbf{x}^0 -al nem azonos eredeti csúcspont esetén legalább egy J -beli indexhez tartozó komponensnek nullának kell lennie. Ha $J = \emptyset$, azaz (1.1)-ben $\mathbf{b} = \mathbf{0}$, akkor a $\mathbf{0}$ pont az egyetlen eredeti csúcspont, egyébként a parametrikus vagy a diszjunktív metszést használhatjuk \mathbf{x}^0 -nak mint határolt felületnek a levágására.

Az \mathbf{x}^0 eredeti csúcspontához tartozó bázismegoldás birtokában érdemes megvizsgálni, hogy van-e olyan $j_0 \in I_R^{(2)}$ index, amelyre $\vartheta_{j_0} > 0$ és $i(j_0) \in N$ teljesül. Ekkor \mathbf{x}^{j_0} az \mathbf{x}^0 -nak olyan szomszédos csúcsa P_k -ban, amely szintén eredeti csúcspont. Ha $f(\mathbf{x}^{j_0}) < f(\mathbf{x}^0)$, akkor szimplex iterációval lépünk át az \mathbf{x}^{j_0} csúcspontra. Ezt folytatva eredeti csúcspontoknak egy véges sorozatát nyerhetjük, amely monoton csökkenő célfüggvényérték sorozattal rendelkezik. Ha a célfüggvény értéke eközben f^0 alá kerül, akkor az f^0 értékét is javítani tudjuk. A konvex metszést ezen eredeti csúcspont sorozat utolsó eleméből hajthatjuk végre. Ha ez a csúcspont degenerált, de a sorozatban volt megengedett csúcspont is, akkor a degenerációból adódható nehézségeket elkerülendő visszaléphetünk a nemdegenerált csúcspontra és onnan hajthatjuk végre a konvex metszést.

6. Metszősík módszer

Az (1.1) feladat megoldását célzó első metszősík módszer Tuy-tól származik [18]. Nála a P_0 poliéder korlátossága volt feltéve. A *Tuy-módszer* csak konvex metszéseket alkalmaz, az $\mathbf{x}^0 \in V(P_k)$ csúcspontot úgy változtatva meg, hogy teljesüljön

$$(6.1) \quad f(\mathbf{x}^0) \equiv f(\mathbf{x}^1)$$

az x^0 tetszőleges x^j szomszédos csúcspontja esetén. A (6.1)-nek eleget tevő pontot lokális csúcs-minimumnak nevezik. Lokális csúcs-minimum véges számú szimplex iterációval nyerhető. A *Tuy-módszer* a következő lépések ciklikus ismétlődéséből áll:

1. Az aktuális poliéder egy csúcsának megkeresése. Ha a poliéder üres, akkor leállhatunk.
2. Lokális csúcs-minimum keresése.
3. Konvex metszés előállítása lokális csúcs-minimum pontból és csatolása a feltételi poliédert meghatározó lineáris feltételhez. Menjünk 1.-re.

A leállást az 1. lépés teszi lehetővé. Ekkor az eddig talált legjobb függvényértéket tároló f^0 az (1.1) optimumértéke és azon megengedett pontok, ahol f^0 felvétetik az optimális megoldások.

A *Tuy-módszer* elméleti végeessége vagy konvergenciája nincs bizonyítva. Megjegyezzük azonban, hogy mindeddig nem ismert olyan numerikus példa, amelyre a *Tuy-módszer* ne fejeződött volna be véges számú lépésben. Mindezek ellenére az eddigi [19] és a dolgozat végén közölt numerikus tapasztalatok is ezen alaplódszer relatív lassúságát mutatják más, kombinatorikus eszközökkel kiegészített módszerekkel szemben. Ez elsősorban abból adódik, hogy a metszősík módszer még akkor is folytatódhat, ha az aktuális poliéder már nem tartalmaz eredeti csúcspontot, így f^0 biztosan nem javul a továbbiakban.

A metszősík módszer végeességét MAJTHAY és WHINSTON úgy biztosítja, hogy azt kiegészíti egy eredeti csúcspontot kereső eljárással [9]. Ha ezen eljárás egy eredeti csúcspontot talált, akkor ebből a csúcspontból vagy valamelyik innen kiindulva elérhető eredeti csúcspontból hajtunk végre konvex metszést. Ha a keresési eljárás során a P_0 egy olyan határoló feltételére bukkanunk, amelyen már nincs eredeti csúcspont, akkor ezt a felületet egy felületmetszéssel kizárjuk a további keresésből. A metszősík módszerrel akkor állhatunk le, ha a maradék poliéder üressé válik, illetve ha az eredeti csúcspontot kereső eljárás vagy a felületmetszés végrehajtása közben kiderül, hogy nincs több eredeti csúcspont. Az itt vázolandó metszősík módszer a *Majthay—Whinston-módszer* ötletén alapszik, annak azonban egy javításaként tekinthető, mivel a [9, 15]-ben közölt eredeti csúcspontot kereső, illetve felületet levágó eljárásoknál hatékonyabbakkal dolgozik és a kvázikonkáv függvény korlátos poliéderen történő minimalizálása mellett a nem-korlátos poliéderen alulról korlátos konkáv függvény esetét is kezelni tudja.

1. Az aktuális poliéder egy csúcsának megkeresése. Ha a poliéder üres, akkor leállhatunk.
2. A 2.1. vagy 3.1. algoritmus végrehajtása. Ha kiderül, hogy nincs több eredeti csúcspont, akkor leállhatunk. Ha eredeti csúcspontot nem tartalmazó határoló felületet találunk, akkor menjünk 5.-re, egyébként 3.-ra.
3. Lépkedés monoton csökkenő célfüggvényérték sorozatot adó szomszédos eredeti csúcspontokon.
4. Vagy menjünk 5.-re, vagy állítsunk elő egy konvex metszést az eredeti csúcspontból és menjünk 6.-ra.
5. Felületmetszés előállítása. Ha kiderül, hogy nincs több eredeti csúcspont, akkor leállhatunk.
6. Az előállított metszés csatolása a feltételi poliédert meghatározó lineáris feltételekhez. Menjünk 1.-re.

és $f_j(x_j)$ konkáv volta miatt

$$\eta_0^j > \eta_1^j \dots > \eta_{t_j}^j.$$

Legyen

$$\tilde{f}_j(x_j) = \begin{cases} f_j(x_j), & \text{ha } t_j = 0, \\ f_j(\xi_{t_j}^j) + \eta_{t_j}^j(x_j - \xi_{t_j}^j), & \text{ha } t_j > 0 \end{cases}$$

és

$$\tilde{f}(\mathbf{x}) = \sum_{j=1}^n \tilde{f}_j(x_j).$$

Mivel $P_0 \subset R_+^n$, ezért az $f(\mathbf{x})$ és $\tilde{f}(\mathbf{x})$ függvények legfeljebb egy korlátos halmazon térhetnek el egymástól. Következésképpen $f(\mathbf{x})$ akkor és csak akkor alulról korlátos P_0 -on, ha $\tilde{f}(\mathbf{x})$ is ugyanilyen. Az $\tilde{f}(\mathbf{x})$ függvénynek P_0 -on való alulról korlátossága egy lineáris programozási feladat megoldásával ellenőrizhető, mivel $\tilde{f}(\mathbf{x})$ lineáris függvény.

7.3. Fixköltséges lineáris programozási feladat

Legyen

$$f(\mathbf{x}) = \sum_{j=1}^n f_j(x_j)$$

és

$$f_j(x_j) = c_j x_j + \varphi_j \delta(x_j) \quad (j = 1, \dots, n),$$

ahol

$$\delta(x_j) = \begin{cases} 1, & \text{ha } x_j > 0, \\ 0, & \text{különben} \end{cases}$$

és $\varphi_j \geq 0$. Az ilyen $f(\mathbf{x})$ célfüggvénnyel rendelkező (1.1) feladatot fixköltséges lineáris programozási feladatnak nevezzük [2]. Az $f(\mathbf{x})$ függvény legfeljebb a

$$\sum_{j=1}^n \varphi_j$$

összeggel térhet el az

$$\tilde{f}(\mathbf{x}) = \sum_{j=1}^n c_j x_j$$

lineáris függvénytől, így $f(\mathbf{x})$ akkor és csak akkor alulról korlátos P_0 -on, ha $\tilde{f}(\mathbf{x})$ is ugyanilyen. Mivel $f(\mathbf{x})$ konkáv R_+^n -ben és $P_0 \subset R_+^n$, így ha $f(\mathbf{x})$ alulról korlátos P_0 -on, akkor a véges optimumérték P_0 egy csúcspontján is felvétetik.

Megjegyezzük, hogy a fixköltséges lineáris programozási feladat metszősík módszerrel történő megoldásakor nehézséget jelent, hogy $f(\mathbf{x})$ csak R_+^n -ben konkáv. Ez a következő ötlettel kerülhető ki [2]. Az aktuális P_k poliéder tetszőleges \mathbf{x}^0 csúcsát tekintve előállítható egy $\hat{f}(\mathbf{x})$ szeparábilis, szakaszonként lineáris, R^n -ben konkáv függvény, amely $f(\mathbf{x})$ minoránsa R_+^n -ben, továbbá \mathbf{x}^0 -ban, valamint \mathbf{x}^0 szomszédos csúcspontjaiban ugyanazt az értéket veszi fel, mint $f(\mathbf{x})$. Ha a konvex

7. Speciális feladatok

Ebben a fejezetben néhány olyan feladattal foglalkozunk, ahol az $f(x)$ függvénynek P_0 -on való korlátossága nyilvánvaló, illetve könnyen ellenőrizhető. Abban az esetben, ha P_0 korlátos, megengedhetjük, hogy $f(x)$ a konkáv függvények halmazát magába foglaló kvázikonkáv függvények halmazába tartozzon, mivel az optimumérték akkor is felvétetik P_0 egy csúcspontján és a konvex metszés előállításánál csak $f(x)$ kvázikonkáv voltát használtuk fel.

7.1. Konkáv kvadratikusan feladat

Legyen

$$f(x) = x'Cx + c'x,$$

ahol C $n \times n$ -es negatív szemidefinit szimmetrikus kvadratikusan mátrix és $c \in R^n$. Ekkor $f(x)$ konkáv R^n -ben. A felbontási tétel alapján [11]

$$P_0 = P_0^\Delta + P_0^<$$

ahol P_0^Δ egy korlátos poliéder és

$$P_0^< = \{x | Ax = 0\}.$$

Adott $x \in P_0$ esetén tekintsünk egy $x = x^{(1)} + x^{(2)}$, $x^{(1)} \in P_0^\Delta$, $x^{(2)} \in P_0^<$ felbontást. Ekkor

$$f(x) = f(x^{(1)}) + 2x^{(1)'}Cx^{(2)} + x^{(2)'}Cx^{(2)} + c'x^{(2)}.$$

Könnyen látható, hogy $f(x)$ akkor és csak akkor alulról korlátos P_0 -on, ha tetszőleges $x^{(2)} \in P_0^<$ esetén $Cx^{(2)} = 0$ és $c'x^{(2)} \geq 0$. Ez a feltétel pedig már lineáris programozási eszközökkel ellenőrizhető.

7.2. Konkáv minimalizálási feladat szeparábilis, szakaszonként lineáris célfüggvény esetén

Legyen

$$f(x) = \sum_{j=1}^n f_j(x_j),$$

ahol $f_j(x_j)$ szakaszonként lineáris konkáv függvény R -ben ($j=1, \dots, n$). Jelölje t_j az $f_j(x_j)$ töréspontjainak számát. Ha $t_j > 0$, akkor legyenek

$$\xi_1^j, \dots, \xi_{t_j}^j$$

az $f_j(x_j)$ töréspontjai növekvő sorrendben és

$$\eta_0^j, \eta_1^j, \dots, \eta_{t_j}^j$$

a töréspontok által meghatározott lineáris egyenletek meredekségei. Nyilván

$$\eta_i^j = \frac{f_j(\xi_{i+1}^j) - f_j(\xi_i^j)}{\xi_{i+1}^j - \xi_i^j}, \quad (i = 1, \dots, t_j - 1)$$

metszést $f(\mathbf{x})$ helyett az $\hat{f}(\mathbf{x})$ függvénnyel hajtjuk végre, akkor a metszés csatolásával a P_k poliéderről levágott pontokban $f(\mathbf{x})$ nem vesz fel az eddig talált legjobb értéknél jobb célfüggvény értéket és \mathbf{x}^0 is a levágott pontok között van. Az $\hat{f}(\mathbf{x})$ előállítását [2] ismerteti.

7.4. Lineáris komplementaritási feladat

Tekintsük a lineáris komplementaritási feladatot [8]:

$$(7.1) \quad \begin{aligned} \mathbf{Ax} + \mathbf{w} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \quad \mathbf{w} \geq \mathbf{0} \\ \mathbf{x}'\mathbf{w} &= 0 \end{aligned}$$

ahol \mathbf{A} $m \times m$ -es négyzetes mátrix, \mathbf{b} m -dimenziós vektor és $\mathbf{x}, \mathbf{w} \in R^m$. A (7.1) feladat átalakítható egy konkáv minimalizálási feladattá.

Legyen

$$f(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^m \min \{x_i, w_i\}.$$

Az $f(\mathbf{x})$ függvény konkáv R^n -ben. A

$$(7.2) \quad \begin{aligned} \min f(\mathbf{x}, \mathbf{w}) \\ \mathbf{Ax} + \mathbf{w} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \quad \mathbf{w} \geq \mathbf{0} \end{aligned}$$

feladat célfüggvénye nyilván alulról korlátos a megengedett pontok poliéderén. Könnyen látható, hogy (7.1) akkor és csak akkor megoldható, ha (7.2) optimumértéke nulla és ebben az esetben (7.1) megoldásai és (7.2) optimális megoldásai azonosak.

Ha (7.2)-t a 6. fejezetben vázolt metszősík módszerrel oldjuk meg és az eredeti csúcspontot kereső eljárás során valamelyik (x_i, w_i) párról kiderül, hogy az aktuális korlátoló felületen mindkettő pozitív, akkor ezen a határoló felületen nem érdemes keresni (7.1) megoldását. Tudjuk, hogy a (7.1) tetszőleges megoldását tekintve az x_i és w_i közül valamelyik nulla, ezért egy olyan felület metszési feltétel állítható elő, amely levágja az aktuális határoló felületet, de meghagyja (7.1) azon megoldásait, amelyek a maradék poliéderben vannak [17].

7.5. Bilineáris programozási feladat

A bilineáris programozási feladat a következő alakban adható meg [15]:

$$(7.3) \quad \begin{aligned} \min \mathbf{x}^{(1)'}\mathbf{Cx}^{(2)} + \mathbf{c}^{(1)'}\mathbf{x}^{(1)} + \mathbf{c}^{(2)'}\mathbf{x}^{(2)} \\ \mathbf{A}^{(1)}\mathbf{x}^{(1)} &= \mathbf{b}^{(1)} \\ \mathbf{x}^{(1)} &\geq \mathbf{0} \\ \mathbf{A}^{(2)}\mathbf{x}^{(2)} &= \mathbf{b}^{(2)} \\ \mathbf{x}^{(2)} &\geq \mathbf{0} \end{aligned}$$

ahol C $n^{(1)} \times n^{(2)}$ -es, $A^{(1)}$ $m^{(1)} \times n^{(1)}$ -es, $A^{(2)}$ $m^{(2)} \times n^{(2)}$ -es mátrix, $b^{(1)} \in R^{m^{(1)}}$, $b^{(2)} \in R^{m^{(2)}}$, $x^{(1)}, c^{(1)} \in R^{n^{(1)}}$, $x^{(2)}, c^{(2)} \in R^{n^{(2)}}$. Feltesszük, hogy a

$$P_0^{(i)} = \{x \in R^{n^{(i)}} | A^{(i)}x = b^{(i)}, x \geq 0\} \quad i = 1, 2$$

poliéderek nem üresek és korlátosak. A (7.3) feladat célfüggvénye nem kvázikonkáv, (7.3) megoldása azonban visszavezethető egy konkáv minimalizálási feladat megoldására. Legyen

$$(7.4) \quad f(x^{(1)}) = \min \{x^{(1)'} C x^{(2)} + c^{(1)'} x^{(1)} + c^{(2)'} x^{(2)} | x^{(2)} \in P_0^{(2)}\}.$$

Az $f(x^{(1)})$ függvény konkáv $R^{n^{(1)}}$ -ben [3]. Tekintsük a

$$(7.5) \quad \begin{aligned} \min f(x^{(1)}) \\ A^{(1)}x^{(1)} &= b^{(1)} \\ x^{(1)} &\geq 0 \end{aligned}$$

feladatot. Könnyen látható, hogy $(\hat{x}^{(1)'}, \hat{x}^{(2)'})$ akkor és csak akkor optimális megoldása (7.3)-nak, ha $\hat{x}^{(1)}$ optimális megoldása (7.5)-nek és $\hat{x}^{(2)}$ az $f(\hat{x}^{(1)})$ meghatározásához megoldandó (7.4) feladat optimális megoldása.

8. Számítástechnikai tapasztalatok

A dolgozatban bemutatott metszősík módszerekre számítógépes programok készültek SIMULA programozási nyelven. A programok a Magyar Tudományos Akadémia IBM/3031-es számítógépén futottak. A módszereket 5×10 -es, illetve 10×20 -as feltételi mátrixú feladatokon ellenőriztük. A mátrixok együtthatóit véletlenszerűen állítottuk elő, a célfüggvénynek pedig olyan szigorúan konkáv kvadratikus függvényt választottunk, amely maximumpontja a feltételi poliéderbe esik. Az 1., illetve 2. táblázat 10—10 db 5×10 -es, illetve 10×20 -as feladat futási eredményeinek átlagos jellemzőit mutatja. A táblázatokban

a = a feladatot megoldó módszer sorszáma

1. *Tuy-módszer* [18],
2. *Majthay—Whinston-módszer* [9]
3. *Majthay—Whinston-módszernek* a 6. fejezetben bemutatott módosítása, a 2.1. algoritmus használatával,
4. a *Majthay—Whinston-módszernek* a 6. fejezetben bemutatott módosítása, a 3.1. algoritmus használatával.

b = a futás átlagos CPU ideje másodpercekben;

c = a megoldáshoz szükséges szimplex iterációk átlagos száma;

d = a konvex metszések átlagos száma;

e = a felület metszések átlagos száma.

1. TÁBLÁZAT

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
1	2,71	50,2	9,7	0
2	1,84	49,8	5,7	2,9
3	1,28	40,1	5,4	1,4
4	1,08	38,2	5,4	1,1

2. TÁBLÁZAT

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
1	28,76	542,1	28,7	0
2	22,21	471,2	8,1	31,8
3	19,59	369,8	8,2	25,6
4	10,02	178,5	8,2	12,1

IRODALOM

- [1] CABOT, A. V. and FRANCIS, R. L., "Solving certain nonconvex quadratic minimization problems by ranking the extreme points", *Operations Research* **18** (1970) 82—86.
- [2] FORGÓ, F., *Nemkonvex és diszkrét programozás*. (Közgazdasági és Jogi Könyvkiadó, Budapest, 1978.)
- [3] GAL, T., *Postoptimal Analysis, Parametric Programming and Related Topics* (McGraw-Hill, New York, 1979).
- [4] GLOVER, F., "Convexity cuts and cut search", *Operations Research* **21** (1973) 123—134.
- [5] GLOVER, F., "Polyhedral convexity cuts and negative edge extensions", *Zeitschrift für Operations Research* **18** (1974) 181—186.
- [6] KONNO, H., "A cutting plane algorithm for solving bilinear programs", *Mathematical Programming* **11** (1976) 14—27.
- [7] KOVÁCS, L. B., *Combinatorial Methods of Discrete Programming* (Akadémiai Kiadó, Budapest, 1980).
- [8] LEMKE, C. E., "Bimatrix equilibrium points and mathematical programming", *Management Science* **11** (1965) 681—689.
- [9] MAJTHAY, A. and WHINSTON, A., "Quasi-concave minimization subject to linear constraints", *Discrete Mathematics* **9** (1974) 35—59.
- [10] MATTHEIS, T. H. and RUBIN, D. S., "A survey and comparison of methods for finding all vertices of convex polyhedral sets", *Mathematics of Operations Research* **5** (1980) 167—185.
- [11] PRÉKOPA, A., *Lineáris programozás* (Bolyai János Matematika Társulat, Budapest, 1968).
- [12] ROCKAFELLAR, R. T., *Convex analysis* (Princeton University Press, Princeton, 1970).
- [13] RÖSSLER, M., „Eine Methode zur Berechnung des optimalen Produktionsprogramms bei konkaven Zielfunktion“, *Unternehmensforschung* **15** (1971) 103—111.
- [14] SEBŐ, A., „Metszősíkok a halmazlefedési feladathoz“, *X. Magyar Operációkutatási Konferencia, Debrecen*, 1980.
- [15] SHERALI, H. D. and SHETTY, C. M., "A finitely convergent algorithm for bilinear programming problems using polar and disjunctive face cuts", *Mathematical Programming* **19** (1980) 14—31.
- [16] SHERALI, H. D. and SHETTY, C. M., "Deep cuts in disjunctive programming", *Naval Research Logistics Quarterly* **27** (1980) 453—476.
- [17] SHERALI, H. D. and SHETTY, C. M., *Optimization with Disjunctive Constraints* (Springer, New York, 1980).

- [18] TUY, H., "Concave programming under linear constraints", *Soviet Mathematics* 5 (1964) 1937—1940.
- [19] ZWART, P. B., "Computational aspects of the use of cutting planes in global optimization", in: *Proceedings of the 1971 annual Conference of the ACM* (ACM, 1971) 457—465.

(Beérkezett: 1982. november 25.)

FÜLÖP JÁNOS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1502 BUDAPEST, XI. KENDE U., 13—17.

FINDING ORIGINAL VERTEX AND ITS APPLICATION MINIMIZING A CONCAVE FUNCTION SUBJECT TO LINEAR CONSTRAINTS

J. FÜLÖP

We minimize a concave function over a convex polyhedron which has a vertex. If the function is lower bounded on the polyhedron, then the problem has a finite minimum and takes it at a vertex of the polyhedron as well. We get a finite cutting plane method generating convex cuts from original vertices. This paper presents a modification of the *Majthay—Whinston method* for finding original vertices and applies it solving concave minimization problem.

A NEMLINEÁRIS PROGRAMOZÁSI FELADAT OPTIMALITÁSI FELTÉTELEINEK DIFFERENCIÁLGEOMETRIAI VIZSGÁLATA

RAPCSÁK TAMÁS

Budapest

A dolgozatban a

$$\min f(\mathbf{x})$$

$$\mathbf{x} = \mathbf{x}(\mathbf{u}) \in R^n, \mathbf{u} \in U_k$$

feladattal foglalkozunk, ahol az $\mathbf{x}(\mathbf{u})$ elemi felület *Gauss-féle paraméteres alakban* van adva.

Erre a feladatra a differenciálgeometriai apparátus felhasználásával természetes módon kiterjeszthetők az egyváltozós függvények szélsőérték helyeire vonatkozó optimalitási kritériumok. Ezután az így nyert tételek geometriai tartalmát vizsgáljuk. Megmutatjuk, hogy ezek a tételek a konvex halmazokra vonatkozó elválasztási tételek egyfajta általánosítását is jelentik, illetve e tételek egy speciális esete arra a kérdésre ad választ, hogy a felület vizsgált pontja mikor lesz elliptikus, parabolikus vagy hiperbolikus.

Végül megmutatjuk, hogy az itt megadott optimalitási feltételek az egyenlőség feltételekkel megadott nemlineáris programozási probléma esetében az ismert optimalitási kritériumokat adják. Ehhez egy új és szemléletes regularitási feltételt definiálunk, amelyről megmutatjuk, hogy az itt tárgyalt feladat esetén az eddig ismert legjobb regularitási feltétellel ekvivalens.

1. Bevezetés

A dolgozatban a

$$\min f(\mathbf{x})$$

(1.1)

$$\mathbf{x} = \mathbf{x}(\mathbf{u}) \in R^n, \mathbf{u} \in U_k$$

$$f, x_i \in C^2, i = 1, \dots, n$$

feladattal foglalkozunk, ahol az $\mathbf{x}(\mathbf{u})$ elemi felület *Gauss-féle paraméteres alakban* van adva és k a paraméterter dimenziója.

Erre a feladatra a differenciálgeometriai apparátus felhasználásával természetes módon kiterjeszthetők az egyváltozós függvények szélsőérték helyeire vonatkozó optimalitási kritériumok. A második részben ezt mutatjuk meg, majd a harmadik részben az így nyert optimalitási feltételek geometriai tartalmát vizsgáljuk. Megmutatjuk, hogy ezek a konvex halmazokra vonatkozó elválasztási tételek egyfajta általánosítását is jelentik. E tételek egy speciális esete pedig arra a differenciálgeometriai kérdésre ad választ, hogy a felület vizsgált pontja mikor lesz elliptikus, parabolikus vagy hiperbolikus.

A negyedik részben az optimalitás itt megadott, illetve a jól ismert elsőrendű szükséges feltétele közötti kapcsolatot mutatjuk be. Ehhez egy új és szemléletes

regularitási feltételt definiálunk, amelyről megmutatjuk, hogy a (4.1) feladat esetén az eddig ismert legjobb regularitási feltétellel ekvivalens.

Végül azt bizonyítjuk be, hogy az itt szereplő másodrendű feltételek a (4.1) feladat esetén az ismert másodrendű feltételeket adják.

2. Az optimalitási feltételek

Ebben a részben differenciálgeometriai eredmények felhasználásával megadjuk az első és másodrendű optimalitási feltételeket. A

$$(2.1) \quad \min f(\mathbf{x})$$

$$\mathbf{x} = \mathbf{x}(\mathbf{u}) \in R^n, \mathbf{u} \in U_k$$

feladatot vizsgáljuk. Az $f(\mathbf{x})$ egy az $\mathbf{x}(\mathbf{u})$ felületet tartalmazó nyílt halmazon van értelmezve. Az U_k a k dimenziós paraméterteret vagy annak egy részhalmazát jelenti. Az $\mathbf{x}(\mathbf{u})$ elemi felületen egy $\mathbf{x}_0 = \mathbf{x}(\mathbf{u}_0)$ ponton áthaladó görbe

$$(2.2) \quad \mathbf{x}(t) = \mathbf{x}(\mathbf{u}(t)) = \begin{pmatrix} x_1(u_1(t), \dots, u_k(t)) \\ \vdots \\ x_n(u_1(t), \dots, u_k(t)) \end{pmatrix} t \in I \quad (I \text{ intervallum})$$

alakban adott, ahol $\mathbf{x}_0 = \mathbf{x}(\mathbf{u}(t_0))$, $t_0 \in I$.

Ennek az érintője a t_0 pontban

$$(2.3) \quad \mathbf{v} = \dot{\mathbf{x}}(t_0) = \left. \frac{d\mathbf{x}(t)}{dt} \right|_{t=t_0} = \begin{pmatrix} \frac{dx_1(t)}{dt} \\ \vdots \\ \frac{dx_n(t)}{dt} \end{pmatrix} \Big|_{t=t_0}.$$

Legyen $E_{n \times k}$ az a mátrix, amelynek oszlopai a feltétel felület \mathbf{x}_0 pontbeli érintősíkját kifesztő $\frac{\partial \mathbf{x}(\mathbf{u}_0)}{\partial u_i}$, $i = 1, \dots, k$ paramétervonalérintők, azaz

$$(2.4) \quad E = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \dots & \frac{\partial x_1}{\partial u_k} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial u_1} & \dots & \frac{\partial x_n}{\partial u_k} \end{pmatrix} \Big|_{\mathbf{u}=\mathbf{u}_0},$$

akkor

$$(2.5) \quad \mathbf{v} = E \cdot \dot{\mathbf{u}}_0, \quad \dot{\mathbf{u}} \in (u_1, \dots, u_k)$$

ahol az $\dot{\mathbf{u}}_0$ vektorok az $\mathbf{u}(t)$ görbék t_0 pontbeli érintői, (u_1, \dots, u_k) pedig a paraméterter.

Ha \mathbf{x}_0 a (2.1) feladat lokális minimum pontja, akkor minden \mathbf{x}_0 ponton áthaladó felületi görbe esetén az $f(\mathbf{x}(\mathbf{u}(t)))$, $t \in I$ egyváltozós függvénynek t_0 -ban loká-

lis minimuma van, tehát a következő feltételek teljesülnek:

$$(2.6) \quad \frac{d}{dt} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} = 0,$$

$$(2.7) \quad \frac{d^2}{dt^2} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} \geq 0.$$

Az általánosság megszorítása nélkül feltehetjük, hogy I minden felületi görbe esetén ugyanazt az intervallumot jelenti.

Így egyszerűen belátható az alábbi állítás is.

Ha minden \mathbf{x}_0 ponton áthaladó felületi görbe esetén az $f(\mathbf{x}(\mathbf{u}(t)))$, $t \in I$ egyváltozós függvényre teljesülnek t_0 -ban a szigorú lokális minimalitást biztosító elegendő feltételek, azaz

$$(2.8) \quad \frac{d}{dt} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} = 0,$$

$$\frac{d^2}{dt^2} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} > 0,$$

akkor \mathbf{x}_0 a (2.1) feladat szigorú lokális minimuma.

A később bizonyítandó 2.1., 2.2. lemmákból következik, hogy a (2.8) feltételek teljesülése esetén a t_0 pontnak van olyan $K \subset I$ környezete, amelyben bármely, az \mathbf{x}_0 ponton áthaladó felületi görbe esetén a

$$(2.9) \quad \frac{d^2}{dt^2} f(\mathbf{x}(\mathbf{u}(t))) \geq 0, \quad t \in K \subset I$$

egyenlőtlenségek teljesülnek. Ennek alapján az állítás a következőképpen bizonyítható:

Tegyük fel, hogy \mathbf{x}_0 nem lokális minimum. Ekkor az \mathbf{x}_0 pont bármely környezetét tekintve az $\mathbf{x}(\mathbf{u})$ felületen létezik olyan \mathbf{x} pont, amelyre $f(\mathbf{x}) < f(\mathbf{x}_0)$.

Válasszunk ki egy tetszőleges $\mathbf{x}_1 = \mathbf{x}(t_1)$, $t_1 \in K$ pontot és egy felületi görbét, amelyik áthalad az $\mathbf{x}_0, \mathbf{x}_1$ pontokon. Figyelembe véve, hogy erre a felületi görbére a (2.8) feltételek teljesülnek, illetve azt, hogy

$$f(t_1) = f(t_0) + \frac{d}{dt} f(t_0) (t_1 - t_0) + \frac{d^2}{dt^2} f(t_1) (t_1 - t_0)^2.$$

$$t_1 = t_0 + \theta(t_1 - t_0), \quad 0 < \theta < 1$$

ellentmondásra jutunk.

A későbbiekben (2.6)-ból, illetve (2.7)-ből származtatjuk az optimalitás első-, illetve másodrendű szükséges feltételeit, (2.8)-ból pedig az elegendő feltételeket.

Elvégezve a differenciálásokat azt kapjuk, hogy

$$(2.10) \quad \frac{d}{dt} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} = \nabla f(\mathbf{x}_0) \mathbf{v}, \quad \mathbf{v} \in T$$

$$(2.11) \quad \frac{d^2}{dt^2} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} = \mathbf{v}^T \nabla^2 f(\mathbf{x}_0) \mathbf{v} + \nabla f(\mathbf{x}_0) \ddot{\mathbf{x}}(t_0), \quad \mathbf{v} \in T,$$

ahol T a feltétel felület érintősíkját, a $\nabla f(\mathbf{x}_0)$ sorvektor az $f(\mathbf{x})$ függvény \mathbf{x}_0 pontbeli gradiense, \mathbf{v}^T pedig a \mathbf{v} vektor transzponáltját jelenti.

A (2.10)-ből azonnal következik, hogy az optimalitás elsőrendű feltétele az, hogy a célfüggvény gradiense ortogonális legyen a feltételi felület érintősíkjára.

Számoljuk ki a $\nabla f(\mathbf{x}_0)\ddot{\mathbf{x}}(t)$ értékét.

2.1. LEMMA. Ha a (2.10) teljesül és $\nabla f(\mathbf{x}_0) \neq 0$, akkor tetszőleges \mathbf{x}_0 ponton áthaladó felületi görbénél

$$(2.12) \quad \nabla f(\mathbf{x}_0)\ddot{\mathbf{x}}(t_0) = |\nabla f(\mathbf{x}_0)|\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k).$$

A lemmában szereplő $\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0$ kifejezés a feltételi felületnek a $\nabla f(\mathbf{x}_0)$ irányú normálvektorra vonatkoztatott második alapformája.

Bizonyítás. A vizsgált felületi görbén vezessük be az s ívhossz paramétert. Ez mindig lehetséges, így az $\mathbf{x}(t)$ görbe $\mathbf{x}(t(s)) = \bar{\mathbf{x}}(s)$ előállítását kapjuk. Az ívhossz szerinti differenciálást vesszővel jelölve nyerjük az alábbi összefüggéseket.

$$(2.13) \quad \begin{aligned} \dot{\mathbf{x}}(t) &= \bar{\mathbf{x}}'(s) \cdot \frac{ds}{dt} \\ \ddot{\mathbf{x}}(t) &= \bar{\mathbf{x}}''(s) \cdot \left(\frac{ds}{dt}\right)^2 + \bar{\mathbf{x}}'(s) \frac{d^2s}{dt^2} \\ \nabla f(\mathbf{x}_0)\ddot{\mathbf{x}}(t_0) &= \nabla f(\mathbf{x}_0) \cdot \bar{\mathbf{x}}''(s_0) \cdot \left(\frac{ds}{dt}\right)^2 \Big|_{t=t_0}. \end{aligned}$$

Az utolsó összefüggésben felhasználtuk azt, hogy az $\dot{\mathbf{x}}(t_0)$ és az $\bar{\mathbf{x}}'(s_0)$ iránya megegyezik, valamint azt, hogy a (2.10) feltétel teljesülése miatt $\nabla f(\mathbf{x}_0)\dot{\mathbf{x}}(t_0) = 0$.

Elvégezve a skaláris szorzást és felhasználva azt, hogy egy görbe első görbülete $\kappa(s) = |\bar{\mathbf{x}}''(s)|$ kapjuk, hogy

$$(2.14) \quad \nabla f(\mathbf{x}_0)\ddot{\mathbf{x}}(t_0) = |\nabla f(\mathbf{x}_0)| \cdot \kappa(s_0) \cos(\nabla f(\mathbf{x}_0), \bar{\mathbf{x}}''(s_0)) \left(\frac{ds}{dt}\right)^2 \Big|_{t=t_0}.$$

Jelölje a vizsgált felületi görbe előjeles normálgörbületét $\kappa_n(s)$ és válasszuk a feltétel felület normálvektorának a $\nabla f(\mathbf{x}_0)$ irányú egységvektort. A felületi görbék görbületére vonatkozó eredmények alapján

$$(2.15) \quad \kappa(s_0) \cos(\nabla f(\mathbf{x}_0), \bar{\mathbf{x}}''(s_0)) = \kappa_n(s_0)$$

$$(2.16) \quad \kappa_n(s_0) = \frac{\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0}{\dot{\mathbf{u}}_0^T \mathbf{A}_c \dot{\mathbf{u}}_0}, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k)$$

ahol az $\dot{\mathbf{u}}_0^T \mathbf{A}_c \dot{\mathbf{u}}_0$ kifejezés a feltételi felület első alapformája.

A (2.15), (2.16) összefüggések és $\left(\frac{ds}{dt}\right)^2 \Big|_{t=t_0} = \dot{\mathbf{u}}_0^T \mathbf{A}_c \dot{\mathbf{u}}_0$ alapján

$$(2.17) \quad \nabla f(\mathbf{x}_0)\ddot{\mathbf{x}}(t_0) = |\nabla f(\mathbf{x}_0)|\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0,$$

ami éppen a lemma állítása.

Ennek alapján a (2.11) feltételek helyett a következőket írhatjuk:

$$(2.18) \quad \frac{d^2}{dt^2} f(\mathbf{x}(\mathbf{u}(t)))|_{t=t_0} = \dot{\mathbf{u}}_0^T \mathbf{E}^T \nabla^2 f(\mathbf{x}_0) \mathbf{E} \dot{\mathbf{u}}_0 + |\nabla f(\mathbf{x}_0)| \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k).$$

Itt $\dot{\mathbf{u}}_0^T \mathbf{E}^T \nabla^2 f(\mathbf{x}_0) \mathbf{E} \dot{\mathbf{u}}_0$ a $\mathbf{v}^T \nabla^2 f(\mathbf{x}_0) \mathbf{v}$ kvadratikus forma T -re való megszorítását jelenti.

Vizsgáljuk most azt az esetet, amikor az $f(\mathbf{x})=f(\mathbf{x}_0)$ nívófelület az \mathbf{x}_0 pont egy környezetében $\mathbf{x}_F(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{n-1}$ $(n-1)$ -dimenziós elemi felületként van megadva és az U_{n-1} paraméterter tartalmazza U_k -t. Ez utóbbi feltétel nem jelenti az általánosság megsértését, mivel megengedett paraméter-transzformációval mindig elérhető.

Jelölje T_1 a nívófelület érintősíkját.

2.2. LEMMA. Ha az $f(\mathbf{x})=f(\mathbf{x}_0)$ nívófelület az \mathbf{x}_0 pont egy környezetében $\mathbf{x}_F(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{n-1}$ $(n-1)$ -dimenziós elemi felületként van megadva és $\nabla f(\mathbf{x}_0) \neq 0$, akkor

$$(2.19) \quad \mathbf{v}^T \nabla^2 f(\mathbf{x}_0) \mathbf{v} = -|\nabla f(\mathbf{x}_0)| \hat{\mathbf{u}}_0^T \mathbf{B}_F \hat{\mathbf{u}}_0, \quad \mathbf{v} \in T_1, \quad \hat{\mathbf{u}}_0 \in (u_1, \dots, u_{n-1})$$

ahol az $\hat{\mathbf{u}}_0^T \mathbf{B}_F \hat{\mathbf{u}}_0$ kifejezés a nívófelületnek a $\nabla f(\mathbf{x}_0)$ irányú normálvektorra vonatkoztatott második alapformája.

Bizonyítás. Legyen $\mathbf{x}_F(\mathbf{u}(t))$ egy tetszőleges felületi görbe, ezért

$$(2.20) \quad f(\mathbf{x}_F(\mathbf{u}(t))) = f(\mathbf{x}_0).$$

Differenciáljuk kétszer t szerint a (2.20) azonosságot. Így azt kapjuk, hogy

$$(2.21) \quad \mathbf{v}^T \nabla^2 f(\mathbf{x}_0) \mathbf{v} + \nabla f(\mathbf{x}_0) \ddot{\mathbf{x}}(t_0) = 0, \quad \mathbf{v} \in T_1.$$

Most a 2.1. lemmát felhasználva azonnal adódik az állítás. Megjegyezzük, hogy a 2.2. lemma egy más bizonyítással [31]-ben is szerepel.

Legyen \mathbf{E}_1 a T_1 -et felfeszítő mátrix. Ezért $\forall \mathbf{v} \in T$ esetén $\mathbf{v} = \mathbf{E} \dot{\mathbf{u}}_0 = \mathbf{E}_1 \hat{\mathbf{u}}_0$, azaz $\hat{\mathbf{u}}_0 = \hat{\mathbf{E}} \dot{\mathbf{u}}_0$, ahol $\hat{\mathbf{E}} = (\mathbf{E}_1^T \mathbf{E}_1)^{-1} \mathbf{E}_1^T \mathbf{E}$. Így az alábbi tételt bizonyítottuk.

2.3. TÉTEL. Ha \mathbf{x}_0 a (2.1) feladat lokális minimum pontja és $\nabla f(\mathbf{x}_0) \neq 0$, akkor az optimalitás szükséges feltételei

$$(2.22) \quad \nabla f(\mathbf{x}_0) \mathbf{E} \dot{\mathbf{u}}_0 = 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k),$$

vagy más alakban

$$\nabla f(\mathbf{x}_0) \frac{\partial \mathbf{x}_0}{\partial u_i} = 0, \quad i = 1, \dots, k,$$

$$(2.23) \quad -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}^T \mathbf{B}_F \hat{\mathbf{E}} \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 \geq 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k).$$

2.4. TÉTEL. Ha

$$\nabla f(\mathbf{x}_0) \cdot \frac{\partial \mathbf{x}_0}{\partial u_i} = 0, \quad i = 1, \dots, k \quad \text{és}$$

$$(2.24) \quad -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}^T \mathbf{B}_F \hat{\mathbf{E}} \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_0 \dot{\mathbf{u}}_0 > 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k),$$

akkor \mathbf{x}_0 a (2.1) feladat szigorú lokális minimuma.

Az optimalitás elsőrendű feltételei azt jelentik, hogy a célfüggvény gradiense ortogonális a feltételi felület érintősíkjára, a másodrendű feltételek pedig azt, hogy a $-\hat{\mathbf{E}}^T \mathbf{B}_F \hat{\mathbf{E}} + \mathbf{B}_C$ mátrix pozitív definit. Ez utóbbi feltétel a legkisebb sajátérték pozitivitásával egyenértékű.

A \mathbf{B}_F mátrix meghatározása a gradiens vektor és a *Hesse-mátrix* felhasználásával [31]-ben található.

3. Az optimalitási feltételek geometriai tartalma

Ebben a részben a 2.3., 2.4. tételek geometriai jelentését vizsgáljuk.

Az előzőekben láttuk, hogy az optimalitási feltételek csak az $f(\mathbf{x})=f(\mathbf{x}_0)$ nívófelületre és az $\mathbf{x}(\mathbf{u})$ feltételi felületre jellemző mennyiségeket tartalmaznak. Ez mutatja, hogy itt a két felület egymáshoz való kapcsolatát kell vizsgálni.

Könnyen igazolható, hogy a (2.22) elsőrendű feltételek azt jelentik, hogy az $f(\mathbf{x})=f(\mathbf{x}_0)$ és az $\mathbf{x}(\mathbf{u})$ felületek első rendben érintkeznek az \mathbf{x}_0 pontban, azaz a magasabb dimenziós felület \mathbf{x}_0 -beli érintősíkjá tartalmazza a másik felület érintősíkját.

Ha figyelembe vesszük a másodrendű feltételeket és azt, hogy optimalizálási problémával foglalkozunk, akkor látható, hogy geometriailag az a kérdés, hogy a két felület mikor metszi át egymást a vizsgált pont egy környezetében. Ahhoz, hogy ezt pontosan megfogalmazzuk, szükségünk van az alábbi definíciókra, amelyekhez az elválasztó hipersíkok fogalmának általánosítását használjuk fel.

3.1. DEFINÍCIÓ. Legyen $X_1 \subset R^n$ k_1 -dimenziós, $X_2 \subset R^n$ k_2 -dimenziós nemüres ponthalmaz, $h(\mathbf{x})$ pedig az R^n -ben vagy annak egy részalmazán értelmezett folytonosan differenciálható függvény.

A $h(\mathbf{x})=h(\mathbf{x}_0)$ nívófelületet elválasztó felületnek nevezzük, ha

$$(3.1) \quad X_1 \subset \{\mathbf{x}/h(\mathbf{x}) \leq h(\mathbf{x}_0)\}, \quad X_2 \subset \{\mathbf{x}/h(\mathbf{x}) \geq h(\mathbf{x}_0)\}.$$

3.2. DEFINÍCIÓ. Két felület nem metszi át egymást, ha létezik elválasztó felület.

E definíciók alapján azt mondhatjuk, hogy a 2.3., illetve a 2.4. tételek arra adnak szükséges, illetve elegendő feltételeket, hogy két, közös ponttal rendelkező felület, amelyek közül az egyik nívófelület, mikor nem metszi át egymást a vizsgált pont egy környezetében és mikor helyezkedik el úgy, hogy a $\nabla f(\mathbf{x}_0)$ irányában a feltétel felület minden pontjának a koordinátája nagyobb legyen, mint a nívófelület megfelelő pontjának a koordinátája. Látható, hogy ezekben az esetekben az elválasztó felület maga a nívófelület.

Nézzük meg, mit lehet mondani akkor, ha mind a két felület *Gauss-féle paraméteres alakban* van adva. Legyen tehát $\mathbf{x}_F(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{k_1}$ k_1 -dimenziós, $\mathbf{x}_C(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{k_2}$ ($k_1 \geq k_2$, $U_{k_1} \supset U_{k_2}$) pedig k_2 -dimenziós felület, amelyeknek az \mathbf{u}_0 paraméterű pontjuk közös. Jelölje a felületek érintősíkjait T_F és T_C . Azt mondjuk, hogy a (3.2) feltétel teljesül, ha az $\mathbf{x}_F(\mathbf{u})$ felületnek létezik olyan \mathbf{n}_F normálisa, hogy a $\frac{\partial \mathbf{x}_F(\mathbf{u})}{\partial u_i}$, $i=1, \dots, k_1$, \mathbf{n}_F vektorok által felfeszített altér tartalmazza az $\mathbf{x}_F(\mathbf{u})$ és az $\mathbf{x}_C(\mathbf{u})$ felületet is. Ebben az esetben az előbbi tételek alapján bizonyíthatók az alábbi állítások.

3.3. TÉTEL. Legyen $\mathbf{x}_F(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{k_1}$ k_1 -dimenziós, $\mathbf{x}_c(\mathbf{u}) \in R^n$, $\mathbf{u} \in U_{k_2}$, ($k_1 \cong k_2$, $U_{k_1} \supset U_{k_2}$) pedig k_2 -dimenziós felület, amelyeknek az \mathbf{u}_0 paraméterű pontjuk közös. Ekkor annak szükséges feltétele, hogy a két felület az \mathbf{u}_0 pont egy környezetében ne metsze egymást az, hogy az $\mathbf{x}_F(\mathbf{u})$ felületnek bármely olyan \mathbf{n}_F normálisára, amelyre a (3.2) feltétel érvényben van, vagy a (3.3), vagy a (3.4) feltételek teljesüljenek.

$$(3.3) \quad \begin{aligned} \mathbf{n}_F \mathbf{E}_c \dot{\mathbf{u}}_0 &= 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \\ -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}_c^T \mathbf{B}_F \hat{\mathbf{E}}_c \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 &\geq 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \end{aligned}$$

$$(3.4) \quad \begin{aligned} \mathbf{n}_F \mathbf{E}_c \dot{\mathbf{u}}_0 &= 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \\ -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}_c^T \mathbf{B}_F \hat{\mathbf{E}}_c \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 &\leq 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}). \end{aligned}$$

3.4. TÉTEL. Az előbbi tétel jelöléseit megtartva, annak elegendő feltétele, hogy a két felület az \mathbf{u}_0 pont egy környezetében ne át metsze egymást az, hogy létezzen az $\mathbf{x}_F(\mathbf{u})$ felületnek olyan \mathbf{n}_F normálisa, amelyre a (3.2) feltétel érvényben van és vagy a (3.5), vagy a (3.6) feltételek teljesüljenek.

$$(3.5) \quad \begin{aligned} \mathbf{n}_F \mathbf{E}_c \dot{\mathbf{u}}_0 &= 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \\ -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}_c^T \mathbf{B}_F \hat{\mathbf{E}}_c \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 &> 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \end{aligned}$$

$$(3.6) \quad \begin{aligned} \mathbf{n}_F \mathbf{E}_c \dot{\mathbf{u}}_0 &= 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}) \\ -\dot{\mathbf{u}}_0^T \hat{\mathbf{E}}_c^T \mathbf{B}_F \hat{\mathbf{E}}_c \dot{\mathbf{u}}_0 + \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 &< 0, \quad \dot{\mathbf{u}}_0 \in (u_1, \dots, u_{k_2}). \end{aligned}$$

Megjegyezzük, hogy a 3.3., 3.4. tételek a konvex halmazokra vonatkozó elválasztási tételek (pl. [17], [23]) egyfajta általánosítását is jelentik.

Ha azt a speciális esetet tekintjük, mikor mindkét felület $(n-1)$ -dimenziós és az egyik felület hipersík, akkor arra a kérdésre kapunk választ, hogy egy elemi felület pontjai mikor esnek az érintősík egyik oldalára a vizsgált pont egy környezetében. A klasszikus differenciálgeometriában a felület egy pontját az ottani *Dupin-féle indikátrix* fajtájától függően elliptikus pontnak, hiperbolikus pontnak, illetve parabolikus pontnak nevezzük ([29], [32], [33]). Mivel itt magasabb dimenziós felületekkel foglalkozunk, ezért ezeket a fogalmakat másképpen definiáljuk.

3.5. DEFINÍCIÓ. Az $(n-1)$ -dimenziós felület egy pontja attól függően elliptikus, parabolikus vagy hiperbolikus, hogy a felület pontjai a vizsgált pont egy elég kis környezetében az érintősík egyik oldalára (de az érintősíkra nem), az érintősík egyik oldalára és az érintősíkra vagy az érintősík mindkét oldalára esnek.

Megjegyezzük, hogy a 3.5. definícióban szereplő tulajdonság a klasszikus differenciálgeometriában mint tétel szerepel. Ismert az alábbi tétel is ([29], [32], [33]).

3.6. TÉTEL. Egy 2-dimenziós felület egy pontja attól függően elliptikus, parabolikus vagy hiperbolikus, hogy ott a *Gauss-görbület* értéke nagyobb, egyenlő vagy kisebb mint nulla.

Látható tehát, hogy a 2.3., 2.4., illetve a 3.3., 3.4. tételek az előbbi állítás általánosítását is jelentik.

Kiegészítésképpen közöljük a parabolikus pontok jellemzésére szolgáló alábbi tételt.

3.7. TÉTEL. Az $\mathbf{x}(\mathbf{u}) \in R^n$, $\mathbf{u} \in \bar{U}_k$, felület \mathbf{x}_0 pontja parabolikus, ha van olyan \mathbf{n}_F normális amelyre

$$(3.7) \quad \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 \geq 0 \quad \text{vagy} \quad \dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 \leq 0, \dot{\mathbf{u}}_0 \in (u_1, \dots, u_k) \quad \text{és}$$

(3.8) a \mathbf{B}_c mátrixnak legfeljebb egy nulla sajátértéke van.

Bizonyítás. A 2.3. és 3.3. tételekből következik, hogy (3.6) szükséges feltétel.

A (3.7) és (3.8) feltételek elegendőségét indirekt úton bizonyítjuk. Tegyük fel, hogy a felület átmetszi az érintősíkot. Ebből következik, hogy kiválaszthatunk olyan felületdarabokat, amelyeknek pontjai az érintősík ellentétes oldalaira esnek, de az érintősíkra nem.

Az $\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0$ második alapforma értéke e felület darabok pontjainak megfelelő irányokat tekintve nulla.

Azonban a \mathbf{B}_c mátrixot diagonál alakra hozva, a (3.7) és (3.8) feltételeket felhasználva azt kapjuk, hogy a második alapforma értéke tetszőleges irányt tekintve pozitív (ha $\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 \geq 0$) vagy negatív (ha $\dot{\mathbf{u}}_0^T \mathbf{B}_c \dot{\mathbf{u}}_0 \leq 0$) kivéve a nulla sajátértékhez tartozó sajátvektor irányát.

Ez ellentmondás, így bebizonyítottuk az állítást.

Végül megemlítjük, hogy a matematikai programozás terminológiáját figyelembe véve (lásd pl. [19], [23]) egy $f(\mathbf{x}) = f(\mathbf{x}_0)$, $\mathbf{x} \in R^n$, $(\nabla f(\mathbf{x}_0) \neq 0)$ nivófelület \mathbf{x}_0 pontja akkor és csak akkor elliptikus, illetve parabolikus, ha a függvény a vizsgált pont egy környezetére nézve szigorúan pszeudokonvex vagy szigorúan pszeudokonkáv, illetve pszeudokonvex vagy pszeudokonkáv.

4. Regularitási feltétel

A most következő részben az optimalitás elsőrendű szükséges feltételéből a (2.22)-ből kiindulva bebizonyítjuk a *Lagrange-tételt* egyenlőségfeltételekkel megadott nemlineáris programozási feladat esetén. Ehhez az állításhoz szükség van egy regularitási feltételre. A nemlineáris programozás elméletében számos ilyen feltétel ismert ([4], [19]).

Mi a differenciálgeometriai tárgyalás alapján megadunk egy szemléletes regularitási feltételt, amelyik az itt tárgyalt esetben ekvivalens a *Karush—Kuhn—Tucker-féle regularitási feltétellel*. Megjegyezzük, hogy az itt szereplő regularitási feltétel más formában megfogalmazva megtalálható [11]-ben, ahol mint másodrendű regularitási feltétel szerepel.

A következőkben tehát a

$$(4.1) \quad \begin{aligned} & \min f(\mathbf{x}) \\ & g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m, \quad i = 1, \dots, m \\ & \mathbf{x} \in R^n, \quad f, g_i \in C^2 \end{aligned}$$

feladattal foglalkozunk. Legyen \mathbf{x}_0 a vizsgált pont és

$$(4.2) \quad M = \{\mathbf{y} / \nabla g_i(\mathbf{x}_0) \cdot \mathbf{y} = 0, \quad i = 1, \dots, m\}.$$

Könnyen látható, hogy az előző részben szereplő feltételi felületnek itt a $g(\mathbf{x}) = 0$, $i = 1, \dots, m$ hiperfelületek metszete felel meg. Ezért a továbbiakban feltételezzük,

hogy az x_0 pont egy környezetében a hiperfelületek metszete elemi felületet határoz meg. Jelölje a továbbiakban ennek a felületnek az érintősíkját T .

4.1. DEFINÍCIÓ. Az x_0 pontban a Q regularitási feltétel teljesül, ha $M=T$.

Ismert az az állítás, hogy ha az M halmaz definíciójában szereplő gradienssek lineárisan függetlenek, akkor $M=T$ ([17]).

Most belátjuk, hogy a Q tulajdonság valóban regularitási feltétel.

4.2. TÉTEL. Ha x_0 a (4.1) feladat optimum pontja és $M=T$, akkor léteznek μ_i , $i=1, \dots, m$ értékek, úgy hogy

$$(4.3) \quad \nabla f(x_0) - \sum_{i=1}^m \mu_i \nabla g_i(x_0) = 0.$$

Bizonyítás. Jelölje M^\perp az M ortogonális komplementer alterét. Ez M definíciója alapján nem más, mint a feltételek x_0 pontbeli gradiensei által felfeszített altér. Mivel x_0 a (4.1) feladat optimum pontja, ezért a $\nabla f(x_0)$ ortogonális a T -re, másrészt $M=T$, amiből következik, hogy $\nabla f(x_0) \in M^\perp$, ami éppen az állítás.

4.3. TÉTEL. Ha tetszőleges $f(x)$ függvény esetén, amelyre x_0 a (4.1) feladat optimum pontja, a (4.3) feltételek teljesülnek, akkor $M=T$.

Bizonyítás. Tegyük fel, hogy nem igaz az állítás. Mivel az $M \supset T$, $M^\perp \subset T^\perp$ (M^\perp , T^\perp ortogonális alterek) relációk teljesülnek, ezért az indirekt feltevés miatt, ezek valódi tartalmazást jelentenek.

Válasszuk most meg a $\nabla f(x_0)$ irányt úgy, hogy a $\nabla f(x_0) \in T^\perp$, $\nabla f(x_0) \notin M^\perp$ teljesüljenek. Ehhez megadható olyan függvény, amelynek x_0 a szigorú, lokális optima a megadott egyenlőség feltételek mellett és a gradiense x_0 -ban $\nabla f(x_0)$. (Ilyen függvény pl. a 2.4. tétel segítségével is megadható.)

Mivel a (4.3) feltételek azt jelentik, hogy $\nabla f(x_0) \in M^\perp$, így ellentmondáshoz jutottunk.

A 4.2., 4.3. tételeknek következménye az alábbi állítás.

4.4. TÉTEL. Tetszőleges $f(x)$ függvény esetén, amelyre x_0 a (4.1) feladat optimum pontja, a (4.3) *Lagrange-feltételek* teljesülésének szükséges és elegendő feltétele, hogy $M=T$ legyen.

Ebből a tételből következik, hogy a (4.1) feladat esetén az itt szereplő regularitási feltételnél nem lehet enyhébbet találni.

Ismeretes, hogy az általános nemlineáris programozási feladatra a *Guignard-féle regularitási feltétel* rendelkezik a fenti tulajdonsággal ([4]).

Legyen

$$(4.4) \quad W\{z | z = \lim_{l \rightarrow \infty} \lambda_l(x_l - x_0), x_l \in X, \lambda_l \in R, \lambda_l > 0 \forall l, \lim_{l \rightarrow \infty} x_l = x_0\},$$

ahol X a (4.1) feladat megengedett pontjainak a halmazát jelenti. A *Guignard-féle regularitási feltétel* a következő:

$$(4.5) \quad M = \overline{W},$$

ahol \overline{W} a W lezártját jelenti.

Az előzőekből következik, hogy esetünkben

$$(4.6) \quad T = \overline{W}.$$

5. Másodrendű feltételek

Ebben a részben megmutatjuk, hogy a (2.24) feltételek a (4.1) feladat esetén az ismert másodrendű elegendő feltételeket adják.

5.1. TÉTEL. Ha a (4.1) feladatnál az x_0 pontban a (4.3) feltételek teljesülnek és a

$$(5.1) \quad g_i(x) = 0, \quad i = 1, \dots, m$$

metszettelület

$$(5.2) \quad x_c(u) \in R^n, \quad u \in U_k$$

alakban van adva, akkor

$$(5.3) \quad -|\nabla f(x_0)| \dot{u}_0^T B_c \dot{u}_0 = \sum_{i=1}^m \mu_i v^T \nabla^2 g_i(x_0) v, \quad v \in T, \quad u_0 \in (u_1, \dots, u_k),$$

ahol $\dot{u}_0^T B_c \dot{u}_0$ a metszettelületnek a $\nabla f(x_0)$ irányú egységvektorra vonatkoztatott második alformája.

Bizonyítás. Mivel a vizsgált felületet minden nívófelület tartalmazza, így a 2.2. lemma és a második alapmennyiségek definíciója alapján kapjuk, hogy

$$(5.4) \quad v^T \nabla^2 g_i(x_0) v = -\dot{u}_0^T \nabla g_i(x_0) \cdot \frac{\partial^2 x_c}{\partial^2 u} \bigg|_{u=u_0} \cdot \dot{u}_0, \quad \dot{u}_0 \in (u_1, \dots, u_k), \quad v \in T,$$

ahol

$$\frac{\partial^2 x_c}{\partial^2 u} = \left(\frac{\partial^2 x_{c1}}{\partial^2 u}, \dots, \frac{\partial^2 x_{cn}}{\partial^2 u} \right).$$

Az (5.4) egyenlőségeket μ_i -vel, $\mu_i = 1, \dots, m$ végigszorozva, majd az egyenlőségeket összeadva kapjuk a (4.3) feltételek felhasználásával, hogy

$$(5.5) \quad \sum_{i=1}^m \mu_i v^T \nabla^2 g_i(x_0) v = -\dot{u}_0^T \nabla f(x_0) \cdot \frac{\partial^2 x_c}{\partial^2 u} \bigg|_{u=u_0} \cdot \dot{u}_0, \quad \dot{u}_0 \in (u_1, \dots, u_k), \quad v \in T,$$

ami éppen az állítás.

Mivel a 2.2. lemma alapján

$$(5.6) \quad v^T \nabla^2 f(x_0) v = -|\nabla f(x_0)| \hat{u}_0^T B_F \hat{u}_0, \quad v \in T_1, \quad \hat{u}_0 \in (u_1, \dots, u_{n-1}),$$

így látható, hogy a (2.25) feltételek a (4.1) feladat esetén valóban az ismert másodrendű feltételeket adják.

IRODALOM

- [1] ARROW, K. J. and ENTHOVEN, A. C., "Quasi concave programming", *Econometria* **29** (1961) 778—800.
- [2] AVRIEL, M., *Nonlinear Programming, Analysis and Methods* (Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1976).
- [3] AVRIEL, M., "r-convex functions", *Mathematical Programming* **2** (1972) 309—323.

- [4] BAZARAA, M. S., and SHETTY, C. M., *Foundations of Optimization* (Springer-Verlag, Berlin, Heidelberg, New York, 1976).
- [5] BELLMANN, R., *Introduction to Matrix Analysis* (McGraw-Hill Book Company, New York, Toronto, London, 1960).
- [6] BEN-ISRAEL, A., BEN TAL, A., ZLOBEC, S., "Optimality conditions in convex programming", *Survey of Mathematical Programming*, Vol. 1. (Akadémiai Kiadó, Budapest, 1979, Editor Prékopa, A.).
- [7] COTTLE, R. W., and FERLAND, J. A., "On pseudo-convex functions of nonnegative variables", *Mathematical Programming* 1 (1971) 95—101.
- [8] EISENHART, L. P., *Riemannian Geometry* (Princeton, Princeton University Press, 1964).
- [9] FARKAS, M., „A feltételes szélsőértékről”, *Matematikai Lapok* 24 (1973) 1—2.
- [10] FERLAND, J. A., "Mathematical programming problems with quasi-convex objective functions", *Mathematical Programming* 3 (1972) 296—301.
- [11] FIACCO, A. V., and MCCORMICK, G. P., *Nonlinear Programming, Sequential Unconstrained Minimization Techniques* (Wiley and Sons, New York, 1968).
- [12] GERENCSÉR, L., "On a close relation between quasi-convex and convex functions and related investigations", *Math. Operationsforschung und Statistik* 4 (1973) Heft 3, 201—211.
- [13] GERENCSÉR, L., „Nemlineáris programozási feladatok megoldása szekvenciális módszerekkel”, *MTA SZTAKI Tanulmányok* 49 (1976).
- [14] GREENBERG, H. J. and PIERSKALLA, W. P., "A review of quasiconvex functions", *Operations Research* (1971) 1553—1570.
- [15] KARUSH, W., Minima of functions of several variables with inequalities as side conditions, Dissertation for the degree of master of science, Chicago, Illinois 1939.
- [16] KÉRI, G., "An examination of nonnegativity and quasi-convexity conditions of quadratic forms on the nonnegative orthant", *Studia Scientiarum Mathematicarum Hungaricae* 6 (1971) 193—196.
- [17] LUENBERGER, D. G., *Introduction to Linear and Nonlinear Programming* (Addison—Wesley Publishing Company Inc., 1973).
- [18] MANGASARIAN, O. L., "Pseudo-convex functions", *SIAM Journal on Control* 3 (1965) 281—290.
- [19] MANGASARIAN, O. L., *Nonlinear Programming* (McGraw-Hill Book Company, 1969).
- [20] MANN, H. B., "Quadratic forms with linear constraints", *Amer. Math. Monthly* 50 (1943) 430—433.
- [21] MARTOS, B., "The direct power of adjacent vertex programming methods", *Management Science* 12 (1965) 241—252.
- [22] MARTOS, B., "Quasi-convexity and quasi-monotonicity in nonlinear programming", *Studia Sci. Math. Hungaricae* 2 (1967) 265—273.
- [23] MARTOS, B., *Nonlinear Programming Theory and Methods* (Akadémiai Kiadó, Bp. 1975).
- [24] MEREAU, P., and PAQUET, J., "Second order conditions for pseudo-convex functions", *SIAM J. App. Math.* 27 (1974) 131—137.
- [25] PONSTEIN, J., "Seven kinds of convexity", *SIAM Review* 9 (1967) 115—119.
- [26] PRÉKOPA, A., Sztochasztikus rendszerek optimalizálási problémáiról, Akadémiai doktori disszertáció, Bp. 1970.
- [27] PRÉKOPA, A., Eine Erweiterung der sogenannten Methode der „zulässige Richtungen“ der nichtlinearen Optimierung auf den Fall quasikonkaver Restriktionsfunktionen, *Math. Operationsforschung und Stat.* 4 (1973).
- [28] PRÉKOPA, A., Az optimalizáláselmélet kialakulásának történetéről, *Alk. Mat. Lapok* 4 (1978) 165—191.
- [29] RAPCSÁK, A. és TAMÁSSY L., *Differenciálgeometria* (Tankönyvkiadó, Bp. 1967).
- [30] RAPCSÁK, T., „A SUMT módszer alkalmazása nem konvex programozási feladatok esetén”, *Alk. Mat. Lapok* 2 (1976) 427—437.
- [31] RAPCSÁK, T., „Az optimalitás másodrendű feltételeiről”, *Alk. Math. Lapok* 4 (1978) 109—116.
- [32] SZOLCSANYI, E., *Differenciálgeometria és vektoranalízis* (Tankönyvkiadó, Bp. 1967).
- [33] SZÓKEFALVI-NAGY, GY., GEHÉR, L. és NAGY, P., *Differenciálgeometria* (Műszaki Könyvkiadó, Budapest, 1979).
- [34] ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, Inc. Englewood Cliffs, 1969).
- [35] WETTLING, W., „Über Minimalbedingungen und Newton-Iteration bei nichtlinearen Optimierungsaufgaben“, *International Series of Numerical Mathematics* 15 (1970) 93—99.
- [36] WETTLING, W. and COLLATZ, L., *Optimization Problems* (Springer-Verlag, New York, Heidelberg, Berlin, 1975).

- [37] WOLKOWICZ, H., "Geometry of optimality conditions and constraint qualifications: the convex case", *Math. Programming* 19 (1980) 32—60.

(Beérkezett: 1982. április 27.)

RAPCSÁK TAMÁS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1250 BUDAPEST, ÜRI U. 49.

THE OPTIMALITY CONDITIONS OF THE NONLINEAR PROGRAMMING PROBLEMS AND DIFFERENTIAL GEOMETRY

T. RAPCSÁK

In this paper the following problem:

$$\min f(\mathbf{x}), \quad \mathbf{x} \in R^n$$

$$\mathbf{x} = \mathbf{x}(\mathbf{u}), \quad \mathbf{u} \in U_k$$

(where $\mathbf{x}(\mathbf{u})$, $\mathbf{u} \in U_k$ is a surface in *Gauss form*) will be discussed.

First the optimality conditions of single variable functions are generalized for the above problem on the basis of the differential geometry. It is demonstrated that these theorems have a close relationship to separation theorems.

In the following section it is shown the relation between the new and the well-known first and second-order optimality conditions.

LOGARITMIKUSAN KONKÁV ELOSZLÁSFÜGGVÉNNYEL KORLÁTOZOTT LINEÁRIS PROGRAMOZÁSI FELADAT MEGOLDÁSÁNAK DUÁLIS MEGKÖZELÍTÉSE

KOMÁROMI ÉVA

Budapest

A dolgozatban a

$$\max_{(u,v) \in K} [\min_{F(y) \geq p} uy + vb], \quad K = \{(u, v) | uA_1 + vA_2 = c; u, v \geq 0\}$$

kétlépcsős programozási feladatot mint a

$$\min_{x \in L} cx, \quad L = \{x | F(A_1 x) \geq p, \quad A_2 x \geq b\}$$

valószínűséggel korlátozott programozási feladat „kvázi duálisát” tekintjük, ahol F folytonosan differenciálható, szigorúan logaritmikusan konkáv együttes valószínűségi eloszlásfüggvény. Olyan algoritmust adunk meg, amelyik egyidejűleg mindkét feladatot megoldja.

1. Bevezetés

A dolgozat az alábbi sztochasztikus lineáris programozási feladattól indul ki:

$$\min cx$$

$$(P) \quad \begin{aligned} \mathcal{P}(A_1 x \geq \beta) &\geq p \\ A_2 x &\geq b \end{aligned}$$

Itt A_1 és A_2 $m \times n$, illetve $m_1 \times n$ méretű mátrixok, b és c m_1 , illetve n dimenziós vektorok; β valószínűségi változók vektora, p előírt valószínűség, $0 < p < 1$.

Feltesszük, hogy a β -t alkotó valószínűségi változók együttes eloszlása szigorúan logaritmikusan konkáv, az $F(y) = \mathcal{P}(\beta < y)$ eloszlásfüggvény folytonosan differenciálható.

Eloszlásfüggvények logaritmikus konkávitásának elégséges feltételeire vonatkozó, mindeztől legerősebb eredmények PRÉKOPA ANDRÁSTÓL származnak [4], [6], [7]. A következő tétel módját nyújt arra, hogy a valószínűségi eloszlások elég széles körére, köztük a normális eloszlásra a (P) feladat konvex programozási eszközökkel megoldható legyen.

1.1. TÉTEL. Az egész R^m téren logaritmikusan konkáv f sűrűségfüggvénnyel generált F eloszlásfüggvény az egész R^m téren logaritmikusan konkáv. Ha f egy $\mathcal{D} \subset R^m$ nyílt konvex halmazon pozitív és szigorúan logaritmikusan konkáv, akkor F szigorúan logaritmikusan konkáv a \mathcal{D} halmazon.

A [3], [5], [8] dolgozatokban, ZOUTENDIJK P2 módszerét kiterjesztve, PRÉKOPA ANDRÁS primál típusú algoritmust dolgozott ki a (P) feladatra arra az általánosabb

esetére, amelyben a célfüggvény folytonosan differenciálható konvex függvény. Szintén az általánosabb feladatra vonatkozik MAYER JÁNOS módszere [1], amely a feladat sajátosságait figyelembe véve hatékonyan kombinálja ZOUTENDIJK P1 módszerét és WOLFE redukált gradiens módszerét.

Jelen dolgozatban a (P) feladatot duális oldalról vizsgáljuk és kihasználjuk az F függvény minden lényeges tulajdonságát. A következő szakaszban igazoljuk, hogy a (P) megoldását indokolt az alábbi „kváziduális”-nak nevezhető feladat megoldása révén keresni:

$$(D) \quad \begin{aligned} & \max \left(\min_{F(y) \cong p} (uy) + vb \right) \\ & uA_1 + vA_2 = c \\ & u, v \geq 0. \end{aligned}$$

Megjegyezzük, hogy a (D) feladat önmagában is figyelemre méltó, a kétlépcsős programozás speciális esete.

A 3. szakaszban duál típusú algoritmust mutatunk be a (D) és (P) feladat egyidejű megoldására. A 4. szakaszban az algoritmus konvergenciáját bizonyítjuk és kitérünk az algoritmussal kapcsolatos számítástechnikai tapasztalatokra.

Köszönettel tartozom DANCs ISTVÁNNAK, aki felhívta a figyelmemet a (P) és (D) feladatok közötti duális kapcsolat lehetőségére.

2. A duális és tulajdonságai

Felhasználva az eloszlásfüggvénynek azt a tulajdonságát, hogy minden változójának monoton növekvő függvénye, a (P)-vel ekvivalens következő feladathoz jutunk:

$$(P^*) \quad \begin{aligned} & \min cx \\ & F(y) \cong p \\ & A_1 x \cong y \\ & A_2 x \cong b \end{aligned}$$

Az $F(y) \cong p$ feltételt helyettesítjük az $\ln F(y) \cong \ln p$ feltétellel.

A (P^*) duálisa a következő feladat:

$$(D^*) \quad \begin{aligned} & \max (cx + u(y - A_1 x) + v(b - A_2 x) + \delta(\ln p - \ln F(y))) \\ & uA_1 + vA_2 = c \end{aligned}$$

$$u_i = \delta \frac{\partial \ln F}{\partial y_i}, \quad i = 1, \dots, m$$

$$u, v, \delta \geq 0$$

A (D^*) célfüggvénye az első feltétel figyelembevételével így írható át:

$$uy + vb + \delta(\ln p - \ln F(y))$$

Most kihasználjuk, hogy F logaritmikusan konkáv. Wolfe dualitási tétele szerint ha (P^*) -nak van optimális x^*, y^* megoldása, akkor (D^*) -nak van olyan δ^*, u^*, v^* megoldása, hogy δ^*, u^*, v^* és y^* kielégíti (D^*) feltételeit és

$$cx^* = u^*y^* + v^*b^* + \delta^*(\ln p - \ln F(y^*)).$$

Ez azonban a lineáris programozás dualitási tétele értelmében csak úgy lehet, ha $\delta^*(\ln p - \ln F(y^*)) = 0$.

Ennélfogva δ^*, u^*, y^* kielégíti az alábbi feltételeket:

$$u_i = \delta \frac{\partial \ln F}{\partial y_i}, \quad i = 1, \dots, m; \quad \delta \geq 0;$$

$$\ln F(y) \geq \ln p, \quad \delta(\ln p - \ln F(y)) = 0.$$

A Kuhn—Tucker-tétel értelmében akkor y^* optimális megoldása a

$$\min u^*y$$

$$(S) \quad \ln F(y) \geq \ln p$$

feladatnak.

Ennélfogva ésszerű a (P^*) , illetve (P) feladat megoldását az alábbi feladat megoldása révén keresni:

$$\max \left(\min_{F(y) \geq p} (uy) + vb \right)$$

$$(D) \quad uA_1 + vA_2 = c$$

$$u, v \geq 0.$$

Ez a feladat nem duálisa ugyan a (P^*) -nak, de a (D) és a duális (D^*) optimális megoldásainak halmazai egybeesnek.

Vegyük észre, hogy az (S) feladat megoldása szemipozitív ($u \geq 0$), de nem pozitív u vektorra — az F eloszlásfüggvény voltából következően — úgy képzendő el, hogy $y_i = \infty$ (vagy y_i olyan érték, amelynél a β_i valószínűségi változó eloszlásfüggvénye 1), ha $u_i = 0$, és az optimalizálás a fennmaradó komponensekre értendő, azaz, ahol $u_i > 0$.

Az F eloszlásfüggvényről feltettük, hogy szigorúan logaritmikusan konkáv és folytonosan differenciálható. Ez azt jelenti, hogy minden $u \geq 0$ vektorra az (S) feladatnak pontosan egy y megoldása van, és a megoldás u -tól folytonosan függ. Helyénvaló tehát az (S) feladattal definiált $u \rightarrow y$ függvénykapcsolatot ($u \geq 0$) így jelölni: $y(u)$. A (D) feladat célfüggvényét tehát így írhatjuk át:

$$uy(u) + vb$$

ahol $u \cdot y(u)$ folytonos és jelentésénél fogva szigorúan konkáv függvénye u -nak ($u \geq 0$). Hátravan még a célfüggvény értelmezése az $u=0$ esetben. Ekkor az y -ra csak az $F(y) \geq p$ feltétel vonatkozik. Vegyük észre, hogy ha y az $F(y) \geq p$ feltételt kielégíti és a

$$\max (uy + vb)$$

$$uA_1 + vA_2 = c$$

$$u, v \geq 0$$

lineáris programozási feladat célfüggvénye korlátos, optimális $(\mathbf{u}^0, \mathbf{v}^0)$ megoldásában $\mathbf{u}^0 = \mathbf{0}$, akkor $(\mathbf{u}^0, \mathbf{v}^0)$ a (D) feladatnak optimális megoldása.

2.1. LEMMA. Ha $(\mathbf{u}^1, \mathbf{v}^1)$ és $(\mathbf{u}^2, \mathbf{v}^2)$ a (D) feladat két optimális megoldása, és $\mathbf{u}^2 \geq \mathbf{0}$, akkor $\mathbf{u}^1 = \delta \mathbf{u}^2$ valamely $\delta \geq 0$ számra.

Bizonyítás. Tegyük fel az állítással ellentétben, hogy nincs olyan $\delta \geq 0$ szám, hogy $\mathbf{u}^1 = \delta \mathbf{u}^2$. Tetszőleges $0 < \lambda < 1$ számra a $(\lambda \mathbf{u}^1 + (1-\lambda)\mathbf{u}^2, \lambda \mathbf{v}^1 + (1-\lambda)\mathbf{v}^2)$ a (D) megengedett megoldása. A hozzá tartozó célfüggvényérték pedig, az $\mathbf{u}\mathbf{y}(\mathbf{u})$ függvény szigorúan konkáv voltából következően és az indirekt feltevés szerint:

$$\begin{aligned} & (\lambda \mathbf{u}^1 + (1-\lambda)\mathbf{u}^2) \mathbf{y}(\lambda \mathbf{u}^1 + (1-\lambda)\mathbf{u}^2) + (\lambda \mathbf{v}^1 + (1-\lambda)\mathbf{v}^2) \mathbf{b} > \\ & > \lambda [\mathbf{u}^1 \mathbf{y}(\mathbf{u}^1) + \mathbf{v}^1 \mathbf{b}] + (1-\lambda) [\mathbf{u}^2 \mathbf{y}(\mathbf{u}^2) + \mathbf{v}^2 \mathbf{b}]. \end{aligned}$$

Ez ellentmondásban van azzal a feltevéssel, hogy $(\mathbf{u}^1, \mathbf{v}^1)$ és $(\mathbf{u}^2, \mathbf{v}^2)$ a (D) optimális megoldásai.

2.2. LEMMA. Ha egy (\mathbf{u}, \mathbf{v}) a (D) feladat optimális megoldása ($\mathbf{u} \geq \mathbf{0}$), akkor $\mathbf{u}_j \mathbf{y}(\mathbf{u}) + \mathbf{v}_j \mathbf{b} \leq \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$ a $K = \{(\mathbf{u}, \mathbf{v}) | \mathbf{u}\mathbf{A}_1 + \mathbf{v}\mathbf{A}_2 = \mathbf{c}; \mathbf{u}, \mathbf{v} \geq \mathbf{0}\}$ poliedrikus halmaz valamennyi $(\mathbf{u}_j, \mathbf{v}_j)$ csúcsára. Ha a K halmaz korlátos és $\mathbf{u}_j \mathbf{y}(\mathbf{u}) + \mathbf{v}_j \mathbf{b} \leq \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$ a K halmaz minden $(\mathbf{u}_j, \mathbf{v}_j)$ csúcsára, akkor (\mathbf{u}, \mathbf{v}) optimális megoldása (D)-nek.

Bizonyítás. Tegyük fel először, hogy (\mathbf{u}, \mathbf{v}) optimális és a K halmaz egy $(\mathbf{u}_{j_0}, \mathbf{v}_{j_0})$ csúcsára az $\mathbf{u}_{j_0} \mathbf{y}(\mathbf{u}) + \mathbf{v}_{j_0} \mathbf{b} > \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$. Az $\mathbf{y}(\mathbf{u})$ folytonossága miatt elég kis $\delta > 0$ esetén az $\mathbf{u}_{j_0} \mathbf{y}(\mathbf{u} + \delta \mathbf{u}_{j_0}) + \mathbf{v}_{j_0} \mathbf{b} > \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$ egyenlőtlenség még fennáll. De az $\mathbf{y}(\mathbf{u})$ definíciójából következően

$$\mathbf{u}\mathbf{y}(\mathbf{u} + \delta \mathbf{u}_{j_0}) = \mathbf{u}\mathbf{y} \left(\frac{1}{1+\delta} \mathbf{u} + \frac{\delta}{1+\delta} \mathbf{u}_{j_0} \right) > \mathbf{u}\mathbf{y}(\mathbf{u}),$$

ezért

$$\left(\frac{1}{1+\delta} \mathbf{u} + \frac{\delta}{1+\delta} \mathbf{u}_{j_0} \right) \mathbf{y}(\mathbf{u} + \delta \mathbf{u}_{j_0}) + \left(\frac{1}{1+\delta} \mathbf{v} + \frac{\delta}{1+\delta} \mathbf{v}_{j_0} \right) \mathbf{b} > \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b},$$

ellentmondásban azzal a feltevéssel, hogy (\mathbf{u}, \mathbf{v}) optimális. Tegyük fel most, hogy a K halmaz korlátos. Ekkor $\max_{(\mathbf{u}_j, \mathbf{v}_j) \in K} (\mathbf{u}_j \mathbf{y}(\mathbf{u}) + \mathbf{v}_j \mathbf{b}) \leq \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$ minden $(\mathbf{u}, \mathbf{v}) \in K$ esetén, a (D) feladat definíciója szerint. Abból tehát, hogy $\max_{(\mathbf{u}_j, \mathbf{v}_j) \in K} (\mathbf{u}_j \mathbf{y}(\mathbf{u}) + \mathbf{v}_j \mathbf{b}) \leq \mathbf{u}\mathbf{y}(\mathbf{u}) + \mathbf{v}\mathbf{b}$ következik, hogy (\mathbf{u}, \mathbf{v}) optimális.

3. Az algoritmus

Az algoritmus a (D) és (P) feladat egyidejű megoldására szolgál. Az algoritmusban az $\mathbf{y}(\mathbf{u})$ függvénynek az alábbi két könnyen igazolható tulajdonságát használjuk még ki:

- (i) $\mathbf{u}^1 \mathbf{y}(\lambda \mathbf{u}^1 + \mathbf{u}^2)$ λ -nak szigorúan csökkenő folytonos függvénye minden $\mathbf{u}^1, \mathbf{u}^2 \geq \mathbf{0}$, $\mathbf{u}^1 \neq \mathbf{u}^2$ vektorra, $\lambda > 0$ esetén;
- (ii) $\mathbf{y}(\lambda \mathbf{u}) = \mathbf{y}(\mathbf{u}) \forall \mathbf{u} \geq \mathbf{0}$, $\lambda > 0$ esetén.

A két tulajdonság együtt maga után vonja, hogy $\mathbf{u}^2 \mathbf{y}(\lambda \mathbf{u}^1 + \mathbf{u}^2)$ λ -nak szigorúan növekvő függvénye.

Az algoritmus a (D) feladat olyan megengedett megoldásainak sorozatát állítja elő, amelyekhez tartozó célfüggvényérték szigorúan növekvő.

Kiindulunk az

$$\mathbf{uA}_1 + \mathbf{vA}_2 = \mathbf{c}, \quad \mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}$$

tetszőleges olyan $(\mathbf{u}^0, \mathbf{v}^0)$ megoldásából, amelyre $\mathbf{u}^0 \geq \mathbf{0}$.

A k -adik iteráció ($k=1, 2, \dots$) az alábbi lépésekből áll:

1. *Lépés:* Az (S) megoldása révén meghatározzuk az $\mathbf{y}^{k-1} (\mathbf{u}^{k-1})$ vektort, majd megoldjuk a

$$\max (\mathbf{u} \mathbf{y}^{k-1} (\mathbf{u}^{k-1}) + \mathbf{v} \mathbf{b})$$

$$\mathbf{uA}_1 + \mathbf{vA}_2 = \mathbf{c}; \quad \mathbf{u}, \mathbf{v} \geq \mathbf{0}$$

lineáris programozási feladatot. Ha (\mathbf{u}, \mathbf{v}) optimális megoldás létezik, folytatjuk a 2. lépésnél. Ha a célfüggvény nem korlátos, folytatjuk a 3. lépésnél.

2. *Lépés:*

— Ha $\mathbf{u} = \delta \mathbf{u}^{k-1}$ valamely $\delta \geq 0$ számra, akkor készen vagyunk, (\mathbf{u}, \mathbf{v}) a (D) optimális megoldása, a (P) optimális megoldását pedig megkapjuk, ha a (P*) feladatot az $\mathbf{y} = \mathbf{y}^{k-1} (\mathbf{u}^{k-1})$ vektort rögzítve megoldjuk.

— Máskülönben olyan kis λ számot választunk, $0 < \lambda < 1$, amelyre

$$\mathbf{u} \mathbf{y}(\lambda \mathbf{u} + (1-\lambda) \mathbf{u}^{k-1}) + \mathbf{v} \mathbf{b} > \mathbf{u}^{k-1} \mathbf{y}(\mathbf{u}^{k-1}) + \mathbf{v}^{k-1} \mathbf{b}$$

Ekkor

$$\mathbf{u}^k = \lambda \mathbf{u} + (1-\lambda) \mathbf{u}^{k-1}$$

$$\mathbf{v}^k = \lambda \mathbf{v} + (1-\lambda) \mathbf{v}^{k-1}$$

3. *Lépés:* Megoldjuk a

$$\max (\mathbf{u} \mathbf{y}(\mathbf{u}^k) + \mathbf{v} \mathbf{b})$$

$$\mathbf{uA}_1 + \mathbf{vA}_2 = \mathbf{0}; \quad \mathbf{u}, \mathbf{v} \geq \mathbf{0};$$

$$u_i \leq 1, \quad i = 1, \dots, m$$

feladatot.

— Ha a célfüggvény nem korlátos, akkor megállapítjuk, hogy a (D) feladatnak nincs optimális megoldása, a (P) feladatnak pedig nincs megengedett megoldása, mert az $\{\mathbf{x} | \mathbf{A}_2 \mathbf{x} \leq \mathbf{b}\}$ halmaz üres. Az eljárás véget ért.

— Máskülönben az optimális (\mathbf{u}, \mathbf{v}) megoldáshoz meghatározzuk az (S) feladat megoldása révén az $\mathbf{y}(\mathbf{u})$ vektort.

Ha $\mathbf{u} \mathbf{y}(\mathbf{u}) + \mathbf{v} \mathbf{b} > 0$, akkor megállapítjuk, hogy a (D) feladat célfüggvénye nem korlátos, nincs tehát optimális megoldása, a (P) feladatnak pedig nincs megengedett megoldása.

Ha $\mathbf{u} \mathbf{y}(\mathbf{u}) + \mathbf{v} \mathbf{b} = 0$, akkor vagy (P*) megoldható az $\mathbf{y} = \mathbf{y}(\mathbf{u})$ vektorra, vagy (P)-nek nincs megengedett megoldása. Az eljárás véget ért.

Ha $\mathbf{u} \mathbf{y}(\mathbf{u}) + \mathbf{v} \mathbf{b} < 0$, akkor olyan λ értéket határozunk meg, amelyre

$$\mathbf{u} \mathbf{y}(\lambda \mathbf{u} + \mathbf{u}^{k-1}) + \mathbf{v} \mathbf{b} \leq 0, \quad \lambda > 0,$$

és

$$\begin{aligned}
 & (\mathbf{u}^{k-1} + \lambda \mathbf{u}) \mathbf{y}(\mathbf{u}^{k-1} + \lambda \mathbf{u}) + (\mathbf{v}^{k-1} + \lambda \mathbf{v}) \mathbf{b} > \\
 & > \mathbf{u}^{k-1} \mathbf{y}(\mathbf{u}^{k-1}) + \mathbf{v}^{k-1} \mathbf{b}
 \end{aligned}$$

Ekkor

$$\mathbf{u}^k = \mathbf{u}^{k-1} + \lambda \mathbf{u}$$

$$\mathbf{v}^k = \mathbf{v}^{k-1} + \lambda \mathbf{v}$$

4. Az algoritmus tulajdonságai

4.1. LEMMA. Tegyük fel, hogy a $K = \{(\mathbf{u}, \mathbf{v}) | \mathbf{u} \mathbf{A}_1 + \mathbf{v} \mathbf{A}_2 = \mathbf{c}; \mathbf{u}, \mathbf{v} \geq 0\}$ halmaz korlátos. Ekkor vagy az algoritmus véges számú iteráció után befejeződik, vagy az algoritmus által generált $(\mathbf{u}^k, \mathbf{v}^k)$ sorozat a (D) feladat olyan $(\mathbf{u}^*, \mathbf{v}^*)$ optimális megoldásához konvergál, amelyben $(\mathbf{u}^* \geq 0)$.

Bizonyítás. A K halmaz korlátossága miatt elemeinek minden végtelen sorozatból kiválasztható konvergens részsorozat. Tegyük fel, hogy az $(\mathbf{u}^k, \mathbf{v}^k)$ sorozat $(\mathbf{u}^k, \mathbf{v}^k)$ részsorozata $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ -hoz konvergál. Igazolni fogjuk, hogy $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ a (D) optimális megoldása. Először belátjuk, hogy $\hat{\mathbf{u}} \geq 0$. Az $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ a K poliéder csúcsainak konvex kombinációja:

$$(\hat{\mathbf{u}}, \hat{\mathbf{v}}) = \sum_{j=1}^s \xi_j (\mathbf{u}_j, \mathbf{v}_j)$$

ahol $(\mathbf{u}_j, \mathbf{v}_j)$ ($j=1, \dots, s$) jelöli a K csúcsait, $\xi_j \geq 0$, $\sum_{j=1}^s \xi_j = 1$. Vegyük észre, hogy ha a K poliéder egy $(\mathbf{u}_r, \mathbf{v}_r)$ csúcsa olyan, hogy $\mathbf{u}_r = 0$, akkor $\xi_r = 0$. Ez abból következik, hogy az $(\mathbf{u}^k, \mathbf{v}^k)$ részsorozat elemeinek a csúcsok konvex kombinációjaként való $\sum_{j=1}^s \xi_j^k (\mathbf{u}_j, \mathbf{v}_j)$ ($\xi_j^k \geq 0$, $\sum_{j=1}^s \xi_j^k = 1$) előállításában sem vehet részt az $(\mathbf{u}_r, \mathbf{v}_r)$ csúcs, különben az algoritmus 2. lépésében az eljárás befejeződött volna. Ezért $\hat{\mathbf{u}} \geq 0$.

Most belátjuk, hogy az $(\mathbf{u}^k, \mathbf{v}^k)$ sorozat $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ -hoz konvergál. Minthogy $\mathbf{u}^k \mathbf{y}(\mathbf{u}^k) + \mathbf{v}^k \mathbf{b}$ szigorúan növekvő, felülről korlátos sorozat, ezért

$$(a) \quad \lim_{k \rightarrow \infty} [\mathbf{u}^k \mathbf{y}(\mathbf{u}^k) + \mathbf{v}^k \mathbf{b}] = \lim_{i \rightarrow \infty} [\mathbf{u}^i \mathbf{y}(\mathbf{u}^i) + \mathbf{v}^i \mathbf{b}] = \hat{\mathbf{u}} \mathbf{y}(\hat{\mathbf{u}}) + \hat{\mathbf{v}} \mathbf{b}.$$

Most vizsgáljuk meg, hogyan konstruálja az algoritmus az $(\mathbf{u}^{k_i+1}, \mathbf{v}^{k_i+1})$ vektort. Minthogy a K halmaz korlátos, az algoritmus 3. lépésére nem kerül sor. Az 1. lépésben meghatározza a K halmaznak azt a csúcsát, legyen ez $(\mathbf{u}_{j_i}, \mathbf{v}_{j_i})$, amelyre $\mathbf{u}_{j_i} \mathbf{y}(\mathbf{u}^{k_i}) + \mathbf{v}_{j_i} \mathbf{b} = \max_{(\mathbf{u}, \mathbf{v}) \in K} [\mathbf{u} \mathbf{y}(\mathbf{u}^{k_i}) + \mathbf{v} \mathbf{b}]$, majd olyan $0 < \lambda = \lambda_{k_i} < 1$ számot, amelyre

$$(b) \quad \mathbf{u}_{j_i} \mathbf{y}(\mathbf{u}^{k_i} + \frac{\lambda}{(1-\lambda)} \mathbf{u}_{j_i}) + \mathbf{v}_{j_i} \mathbf{b} > \mathbf{u}^{k_i} \mathbf{y}(\mathbf{u}^{k_i}) + \mathbf{v}^{k_i} \mathbf{b}$$

Az $(\mathbf{u}^{k_i+1}, \mathbf{v}^{k_i+1})$ vektor, ahol $\mathbf{u}^{k_i+1} = \lambda \mathbf{u}_{j_i} + (1-\lambda) \mathbf{u}^{k_i}$ és $\mathbf{v}^{k_i+1} = \lambda \mathbf{v}_{j_i} + (1-\lambda) \mathbf{v}^{k_i}$, az alábbi egyenlőtlenségeket kielégíti:

$$\mathbf{u}^k \mathbf{y}(\mathbf{u}^{k_i}) + \mathbf{v}^k \mathbf{b} < \mathbf{u}^{k_i+1} \mathbf{y}(\mathbf{u}^{k_i+1}) + \mathbf{v}^{k_i+1} \mathbf{b} < \hat{\mathbf{u}} \mathbf{y}(\hat{\mathbf{u}}) + \hat{\mathbf{v}} \mathbf{b}.$$

Tovább ez azt jelenti, hogy

$$\begin{aligned} 0 &< (1 - \lambda_{k_i}) \mathbf{u}^{k_i} [\mathbf{y}(\mathbf{u}^{k_i+1}) - \mathbf{y}(\mathbf{u}^{k_i})] \\ &+ \lambda_{k_i} [\mathbf{u}_{j_i} \mathbf{y}(\mathbf{u}^{k_i+1}) + \mathbf{v}_{j_i} \mathbf{b} - (\mathbf{u}^{k_i} \mathbf{y}(\mathbf{u}^{k_i}) + \mathbf{v}^{k_i} \mathbf{b})] < \\ &< \hat{\mathbf{u}} \mathbf{y}(\hat{\mathbf{u}}) + \hat{\mathbf{v}} \mathbf{b} - \sum_{j=1}^s \xi_j^{k_i} \left[\mathbf{u}_j \mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j \right) + \mathbf{v}_j \mathbf{b} \right] = \varepsilon_{k_i} \end{aligned}$$

(a)-ból következik, hogy $\varepsilon_{k_i} \rightarrow 0$, ha $i \rightarrow \infty$, (b)-ből pedig az, hogy két pozitív tag összege kisebb, mint ε_{k_i} . Ez azt jelenti, hogy egyenként 0-hoz tartanak:

$$0 < (1 - \lambda_{k_i}) \sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j \left[\mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j + \frac{\lambda_{k_i}}{1 - \lambda_{k_i}} \mathbf{u}_{j_i} \right) - \mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j \right) \right] \rightarrow 0,$$

$$0 < \lambda_{k_i} \left(\mathbf{u}_{j_i} \mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j + \frac{\lambda_{k_i}}{1 - \lambda_{k_i}} \mathbf{u}_{j_i} \right) + \mathbf{v}_{j_i} \mathbf{b} - \sum_{j=1}^s \xi_j^{k_i} \left[\mathbf{u}_j \mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j \right) + \mathbf{v}_j \mathbf{b} \right] \right) \rightarrow 0,$$

ha $i \rightarrow \infty$.

Figyelembe véve azt, hogy $\xi_j^{k_i} \rightarrow \xi_j$ ($j=1, \dots, s$), ez implikálja, hogy

$$\lambda_{k_i} \rightarrow 0$$

és

$$(\mathbf{u}^{k_i+1}, \mathbf{v}^{k_i+1}) \rightarrow (\hat{\mathbf{u}}, \hat{\mathbf{v}}),$$

ha $i \rightarrow \infty$. Ezzel beláttuk, hogy az $(\mathbf{u}^k, \mathbf{v}^k)$ sorozat $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ -hoz konvergál.

Végül igazoljuk, hogy $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ optimális. A

$$\lim_{i \rightarrow \infty} \left(\mathbf{u}_{j_i} \mathbf{y} \left(\sum_{j=1}^s \xi_j^{k_i} \mathbf{u}_j + \frac{\lambda_{k_i}}{1 - \lambda_{k_i}} \mathbf{u}_{j_i} \right) + \mathbf{v}_{j_i} \mathbf{b} \right) = \sum_{j=1}^s \xi_j \left(\mathbf{u}_j \mathbf{y} \left(\sum_{j=1}^s \xi_j \mathbf{u}_j \right) + \mathbf{v}_j \mathbf{b} \right)$$

összefüggésből következik, hogy

$$\max_{(\mathbf{u}_j, \mathbf{v}_j) \in K} (\mathbf{u}_j \mathbf{y}(\hat{\mathbf{u}}) + \mathbf{v}_j \mathbf{b}) \leq \hat{\mathbf{u}} \mathbf{y}(\hat{\mathbf{u}}) + \hat{\mathbf{v}} \mathbf{b},$$

amely a 2.2. lemma szerint bizonyítja, hogy $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ optimális. Ezzel a 4.1. lemma bizonyítását befejeztük.

5. Számítástechnikai megjegyzések

Az algoritmust olyan példákon vizsgáltam, amelyekben az F eloszlásfüggvény normális eloszlású független valószínűségi változók együttes eloszlásfüggvénye. Az algoritmus 2. lépésében szereplő λ értékét előre rögzítettem, vagyis azt a tulajdonságot, hogy

$$\mathbf{u} \mathbf{y}(\lambda \mathbf{u} + (1 - \lambda) \mathbf{u}^{k-1}) + \mathbf{v} \mathbf{b} > \mathbf{u}^{k-1} \mathbf{y}(\mathbf{u}^{k-1}) + \mathbf{v}^k \mathbf{b}$$

nem követeltem meg. A 3. lépésben pedig λ kiválasztásánál csak arra ügyeltem, hogy az $\mathbf{u} \mathbf{y}(\lambda \mathbf{u} + \mathbf{u}^{k-1}) + \mathbf{b} \leq 0$ egyenlőtlenség fennálljon. Ezek miatt néhány példában a (D) feladat célfüggvényértéke átmenetileg csökkent.

A megengedett megoldások K halmaza a példák egyikében sem volt korlátos. A lineáris feltételrendszer $(\mathbf{A}'_1, \mathbf{A}'_2)$ mátrixának méretei 2×8 , 17×19 , 21×31 ,

28×36 és 28×53 voltak. A 2×8 méretű próbafeladaton kívül a többiek a [8] dolgozatban leírt STABIL sztochasztikus programozási modell adataiból származnak. A STABIL modell c vektora mellett még két másik vektorral is számoltam. A példákban a sztochasztikus változók m száma (az u vektor mérete) 3 és 8 között változott.

A konvergencia gyorsaságára vonatkozó első számítástechnikai tapasztalatok kedvezőek. Legfeljebb 5 iteráció után a megoldás mindegyik feladatnál a 2%-os relatív pontosságon belül került. Az 5 iterációban 7 LP feladat megoldása szerepelt és 15 alkalommal az (S) feladat megoldása. Az a tapasztalat, hogy az iterációk és bennük az (S) feladatok száma a sztochasztikus változók számával nő.

Az algoritmus vizsgálata a Számítástechnikai és Automatizálási Kutató Intézet IBM 3031 számítógépén történt. A lehetőségért az Intézetnek köszönettel tartozom.

IRODALOM

- [1] MAYER, J., "A nonlinear programming method for the solution of a stochastic programming model of A. Prékopa", *Survey of Math. Programming, Proc. of the 9th Int. Math. Progr. Symp.*, ed. A. Prékopa, Vol. 2 (1979) 129—139.
- [2] MILLER, L. B. and WAGNER, H., "Chance-constrained programming with joint constraints", *Operations Research* 13 (1965) 930—945.
- [3] PRÉKOPA, A., "On probabilistic constrained programming", *Proc. of the Princeton Symp. on Math. Programming* (1970) 113—138.
- [4] PRÉKOPA, A., "Logarithmic concave measures with application to stochastic programming", *Acta Sci. Math.* 32 (1971) 301—316.
- [5] PRÉKOPA, A., "A megengedett irányok elnevezésű nemlineáris programozási módszer kiterjesztése kvázikonkáv feltételi függvények esetére", *MTA Számítástechnikai Központ Közleményei* 9 (1972) 3—16.
- [6] PRÉKOPA, A., "On logarithmic concave measures and functions", *Acta Sci. Math.* 34 (1973) 335—343.
- [7] PRÉKOPA, A., "Contributions to the theory of stochastic programming", *Math. Programming* 4 (1973) 202—221.
- [8] PRÉKOPA, A., GANCZER, S., DEÁK, I. és PATYI, K., "A STABIL sztochasztikus programozási modell és annak kísérleti alkalmazása a magyar villamosenergia-iparra", *Alk. Mat. Lapok* 1 (1975) 3—22.

(Beérkezett: 1982. február 19.)

KOMÁROMI ÉVA
ORSZÁGOS VEZETŐKÉPZŐ KÖZPONT
1087 BUDAPEST, KÖNYVES K. KRT. 48—52.

A DUAL APPROACH TO STOCHASTIC LINEAR PROGRAMMING PROBLEMS CONSTRAINED BY LOGARITHMIC CONCAVE JOINT PROBABILITY DISTRIBUTION FUNCTION

É. KOMÁROMI

This paper considers the following two-stage programming problem: $\max_{(u,v) \in K} [\min_{x \in L} uy + vb]$, $K = \{(u, v) | uA_1 + vA_2 = c, u, v \geq 0\}$ as a "quasidual" to the chance-constrained programming problem: $\min_{x \in L} cx$, $L = \{x | F(A_1x) \geq p, A_2x \leq b\}$, where F is a continuously differentiable strictly logarithmic concave joint probability distribution function. An algorithm is presented for solving both problems simultaneously.

A CSEBISEV-EGYENLŐTLENSÉG EGY ÉLESÍTÉSE MONTE CARLO-MÓDSZERREL BECSÜLT FÜGGVÉNYÉRTÉKEK ESETÉN

PINTÉR JÁNOS

Budapest

A *Csebisev-egyenlőtlenség* egy élesítését bizonyítjuk és vetjük össze a *Bernstein-féle élesítéssel*, relatív gyakoriság alapján becsült valószínűségek esetén. A javítás mértékét vizsgáljuk és szám-példákkal illusztráljuk. Végül a vizsgált módszernek zajjal torzított függvényértékek kiszámítására vonatkozó kiterjesztését ismertetjük.

1. Bevezetés

Sztochasztikus optimalizálási feladatok megoldása során a feladatban szereplő függvények értékét gyakran csak közelítőleg tudjuk kiszámítani, a függvényértékeket befolyásoló véletlen tényezők nagyszámú realizációja alapján (lásd pl. [1, 4, 5]). Természetes követelmény, hogy ilyen esetekben a véletlen tényezők realizációi számát lehetőleg csökkentsük (adott pontosságú és megbízhatóságú becslés biztosítása mellett), hiszen a függvényérték-becslésre fordított erőforrások (kísérletek, számítógépi szimulációs ciklusok) mennyisége igen jelentős lehet.

Az ismeretlen valószínűségek becsléséhez szükséges kísérletek (realizációk) számát gyakran az általánosan ismert *Csebisev-egyenlőtlenség* alapján szokták meghatározni. Eszerint egy tetszőleges $0 < p < 1$ valószínűség $0 < \varepsilon \leq pq$ pontosságú és $0 < 1 - \delta < 1$ megbízhatóságú becsléséhez szükséges független kísérletek számát az

$$n \cong \frac{1}{4\delta\varepsilon^2} \cong \frac{pq}{\delta\varepsilon^2} \quad (q = 1 - p)$$

egyenlőtlenségből lehet meghatározni (lásd pl. [2], 309. old.). BERNSTEIN megmutatta ([2], 319—323. old., [3]), hogy ez az eredmény lényegesen javítható, és már

$$n \cong \ln \frac{2}{\delta} \cdot \frac{9}{8\varepsilon^2} \cong \ln \frac{2}{\delta} \cdot \frac{2pq \left(1 + \frac{\varepsilon}{2pq}\right)^2}{\varepsilon^2}$$

számú realizáció is elégséges.

Szokásos még a realizációk számának a centrális határeloszlás-tétel alapján történő becslése. Itt kihasználják, hogy a relatív gyakoriság standardizáltja közelítőleg normális eloszlást követ; ennek alapján a realizációk számára adódó becslés

$$n \cong \frac{k^2(\delta)}{4\varepsilon^2} \cong \frac{pqk^2(\delta)}{\varepsilon^2},$$

ahol $k(\delta)$ értéke az

$$\frac{1}{\sqrt{2\pi}} \int_{-k(\delta)}^{k(\delta)} e^{-\frac{x^2}{2}} dx = 1 - \delta$$

egyenletből határozható meg (lásd pl. [6], 78. old.). A fenti aszimptotikus becsléssel kapcsolatban megjegyezzük, hogy — az n számú független kísérlet alapján adódó relatív gyakoriság standardizáltjának eloszlásfüggvényét $F_n(x)$ -szel, a standard normális eloszlásfüggvényt pedig $\Phi(x)$ -szel jelölve — a *Berry—Esseen-tétel* szerint (lásd pl. [6], 75. old.) fennáll

$$\sup_{-\infty \leq x \leq \infty} |F_n(x) - \Phi(x)| \leq \frac{p^2 + q^2}{\sqrt{npq}},$$

és a becslés $O\left(\frac{1}{\sqrt{npq}}\right)$ nagyságrendje nem javítható. Ez azt jelenti, hogy az $F_n(x) \sim \Phi(x)$ approximáció 0-hoz (1-hez) közeli p valószínűségekre még nagy n értékek esetén is elég pontatlan lehet. A sztochasztikus programozási feladatokban viszont általában éppen ilyen típusú valószínűségek becslésére van szükség, vagyis n értékének fenti aszimptotikus közelítése gyakran nem megengedett.

Dolgozatunkban a *Csebishev-egyenlőtlenség* egy élesítését bizonyítjuk, relatív gyakoriság alapján becsült valószínűségek esetén. Eredményünket a *Bernstein-féle élesítéssel* hasonlítjuk össze, a javítás mértékét speciális esetekben vizsgáljuk és számpéldákkal illusztráljuk. Végül a vizsgált módszernek zajjal torzított függvényértékek kiszámítására vonatkozó kiterjesztését ismertetjük.

2. A Csebishev-egyenlőtlenség élesítése relatív gyakoriság alapján becsült valószínűségek esetén

Legyen $0 < p < 1$ tetszőleges valószínűség, ξ_k^* , $k=1, 2, 3, \dots$ pedig p paraméterű *Bernoulli-eloszlású*, teljesen független valószínűségi változók sorozata, azaz

$$(2.1) \quad P(\xi_k^* = 1) = p, \quad P(\xi_k^* = 0) = q = 1 - p, \quad k = 1, 2, 3, \dots$$

Szimmetria okokból feltehetjük, hogy (2.1)-ben $p \leq \frac{1}{2}$.

Legyen $\xi_k = \xi_k^* - M(\xi_k^*) = \xi_k^* - p$, $k=1, 2, 3, \dots$, ekkor természetesen $M(\xi_k) = 0$, $D(\xi_k) = \sqrt{pq}$. Tetszőleges $v > 0$ esetén tekintsük az $e^{v\xi_k}$ valószínűségi változót.

2.1. LEMMA.

$$(2.2) \quad M(e^{v\xi_k}) = e^{-vp}(1 - p + pe^v).$$

Bizonyítás. A ξ_k valószínűségi változók definíciója alapján elemi úton adódik a (2.2) reláció:

$$M(e^{v\xi_k}) = pe^{v(1-p)} + (1-p)e^{v(-p)} = e^{-vp}(1 - p + pe^v).$$

2.2. LEMMA. Tetszőleges $v > 0$ és $p \leq \frac{1}{2}$ esetén

$$(2.3) \quad M(e^{v\xi_k}) \leq M(e^{-v\xi_k}).$$

Bizonyítás. A (2.2) reláció felhasználásával az

$$(2.4) \quad (1-p)e^{-vp} + pe^{v(1-p)} \cong (1-p)e^{vp} + pe^{-v(1-p)}$$

egyenlőtlenséget kell igazolnunk. Ez nyilván ekvivalens a

$$G(p) = \frac{e^{vp} - e^{-vp}}{p} \cong \frac{e^{v(1-p)} - e^{-v(1-p)}}{1-p} = G(1-p)$$

egyenlőtlenséggel. Itt viszont

$$G(p) = \frac{2\text{sh}(vp)}{p} = 2 \sum_{j=0}^{\infty} \frac{v^{2j+1} p^{2j}}{(2j+1)!}$$

p -nek monoton növekvő függvénye. Ezért a $p \cong \frac{1}{2}$ feltevés értelmében $G(p) \cong G\left(\frac{1}{2}\right) \cong G(1-p)$, vagyis fennáll a (2.4) reláció.

Vezessük be a

$$\xi^{*(n)} = \sum_{k=1}^n \xi_k^*, \quad \xi^{(n)} = \sum_{k=1}^n \xi_k,$$

$$\mathcal{M}(v) = M(e^{v\xi^{(n)}}), \quad \mathcal{N}(v) = M(e^{-v\xi^{(n)}}) \quad (v > 0)$$

jelöléseket, ekkor $M\left(\frac{\xi^{*(n)}}{n}\right) = p$, $M(\xi^{(n)}) = 0$.

2.3. LEMMA. Legyen ε tetszőleges pozitív szám, ekkor

$$(2.5) \quad P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \geq \varepsilon\right) \leq 2e^{-n[v(\varepsilon+p) - \ln(1-p+pe^v)]}.$$

Bizonyítás. A Csebisev-egyenlőtlenség Bernstein-féle élesítésénél alkalmazott technikát használjuk fel. Ismeretes (lásd [2], 309–310. old., [3]), hogy a Markov-egyenlőtlenséget valamely

$$\eta = e^{v[\xi - M(\xi)]} \quad (v > 0)$$

valószínűségi változóra alkalmazva, tetszőleges $t > 0$ esetén fennáll a

$$P\left(\xi \geq M(\xi) + \frac{t + \ln M(e^{v[\xi - M(\xi)]})}{v}\right) \leq e^{-t}$$

reláció. Ezt az egyenlőtlenséget a $\xi = \xi^{(n)}$ valószínűségi változóval felírva, $M(\xi^{(n)}) = 0$ miatt adódik

$$P\left(\xi^{(n)} \geq \frac{t + \ln \mathcal{M}(v)}{v}\right) \leq e^{-t},$$

$\xi = -\xi^{(n)}$ esetén pedig

$$P\left(-\xi^{(n)} \geq \frac{t + \ln \mathcal{N}(v)}{v}\right) \leq e^{-t}.$$

A (2.3) reláció értelmében (a ξ_k valószínűségi változók függetlensége miatt) igaz az

$$(2.6) \quad \mathcal{M}(v) = [M(e^{v\xi_k})]^n = [e^{-vp}(1-p+pe^v)]^n \cong [M(e^{-v\xi_k})]^n = \mathcal{N}(v)$$

egyenlőtlenség. Másrészt tetszőleges $t > 0$ esetén

$$\varepsilon = \frac{t + \ln \mathcal{M}(v)}{nv} = \frac{t + n[-vp + \ln(1-p+pe^v)]}{nv} > 0,$$

tehát (2.6) felhasználásával adódik a bizonyítandó (2.5) reláció:

$$\begin{aligned} P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \cong \varepsilon\right) &= P\left(\left|\frac{\xi^{(n)}}{n}\right| \cong \varepsilon\right) \cong P\left(\frac{\xi^{(n)}}{n} \cong \varepsilon\right) + P\left(\frac{-\xi^{(n)}}{n} \cong \varepsilon\right) \cong \\ &\cong P\left(\frac{\xi^{(n)}}{n} \cong \frac{t + \ln \mathcal{M}(v)}{nv}\right) + P\left(\frac{-\xi^{(n)}}{n} \cong \frac{t + \ln \mathcal{N}(v)}{nv}\right) \cong 2e^{-t} = \\ &= e^{-n[v(\varepsilon+p) - \ln(1-p+pe^v)]}. \end{aligned}$$

Tegyük fel, hogy $0 < \varepsilon \leq pq$, és vezessük be az

$$(2.7) \quad F(v, \varepsilon, p) = v(\varepsilon+p) - \ln(1-p+pe^v)$$

jelölést. A következő lépésben a (2.5) egyenlőtlenség élessé tétele érdekében a (2.7) függvény értékét — rögzített ε és p mellett — v szerint maximalizáljuk.

2.4. LEMMA.

$$\max_{v>0} F(v, \varepsilon, p) = F(v_*, \varepsilon, p),$$

ahol

$$(2.8) \quad v_* = \ln \left[\frac{(p+\varepsilon)q}{p(q-\varepsilon)} \right] > 0.$$

Bizonyítás. Az F függvényt v szerint differenciálva a deriváltfüggvény zérushelyét keressük:

$$\frac{\partial F}{\partial v} = \varepsilon + p - \frac{pe^v}{1-p+pe^v} = \frac{(\varepsilon+p)(1-p) + (\varepsilon+p)pe^v - pe^v}{1-p+pe^v} = 0.$$

Itt a pozitív nevező elhagyható, így az

$$(\varepsilon+p)(1-p) + e^vp(\varepsilon+p-1) = 0$$

egyenletet kapjuk, ahonnan (2.8) adódik. Nyilvánvaló, hogy $v < v_*$ esetén $\frac{\partial F}{\partial v} > 0$,

$v > v_*$ esetén pedig $\frac{\partial F}{\partial v} < 0$, v_* tehát valóban az F függvény egyetlen v szerinti maximumhelye.

A 2.3. és 2.4. segédétel felhasználásával közvetlenül adódik a következő eredmény.

2.1. TÉTEL. Tetszőleges $0 < p \leq \frac{1}{2}$ és $0 < \varepsilon \leq pq$ esetén fennáll a

$$(2.9) \quad P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \geq \varepsilon\right) \leq 2e^{-C_*(\varepsilon, p)n}$$

egyenlőtlenség; a $C_*(\varepsilon, p)$ állandó értéke a (2.8)-beli v_* felhasználásával

$$(2.10) \quad C_*(\varepsilon, p) = v_*(\varepsilon + p) - \ln(1 - p + pe^{v_*}) = \ln\left[\left(\frac{\varepsilon + p}{p}\right)^{\varepsilon + p} \left(\frac{1 - p}{1 - p - \varepsilon}\right)^{\varepsilon + p - 1}\right].$$

2.1. *Megjegyzés.* A tétel eredménye hasonló jellegű a *Csebisev-egyenlőtlenség* említett *Bernstein-féle élesítéséhez*. BERNSTEIN tétele teljesen független, korlátos valószínűségi változók esetében alkalmazható, és a (2.1) *Bernoulli-eloszlású* valószínűségi változókkal képezett $\frac{\xi^{*(n)}}{n}$ relatív gyakoriság és a becsülni kívánt p valószínűség eltérésére a

$$(2.11) \quad P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \geq \varepsilon\right) \leq 2e^{-C(\varepsilon, p)n} \quad (0 < \varepsilon \leq pq)$$

egyenlőtlenséget szolgáltatja; itt

$$(2.12) \quad C(\varepsilon, p) = \frac{\varepsilon^2}{2pq\left(1 + \frac{\varepsilon}{2pq}\right)^2}.$$

Ebből következik, hogy a

$$P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \geq \varepsilon\right) \leq \delta \quad (0 < \varepsilon \leq pq \quad 0 < \delta < 1)$$

reláció teljesüléséhez (2.9), illetve (2.11) alapján elegendő

$$(2.13) \quad n \geq n_*(p, \varepsilon, \delta) = \ln \frac{2}{\delta} \frac{1}{C_*(\varepsilon, p)},$$

illetve

$$(2.14) \quad n \geq n(p, \varepsilon, \delta) = \ln \frac{2}{\delta} \frac{1}{C(\varepsilon, p)}.$$

2.2. *Megjegyzés.* A bizonyítás során feltettük, hogy a becsülni kívánt p valószínűsége vonatkozóan teljesül $0 < p \leq \frac{1}{2}$. Ez nem jelent tényleges megszorítást,

hiszen tetszőleges A eseménynek és \bar{A} ellentettjének bekövetkezése egyidejűleg vizsgálható, és

$$P\left(\left|\frac{\xi^{*(n)}}{n} - p\right| \geq \varepsilon\right) \leq \delta$$

ekvivalens a

$$P\left(\left|\frac{n-\xi^{*(n)}}{n}-(1-p)\right|\geq \varepsilon\right) \leq \delta$$

relációval.

A továbbiakban megmutatjuk, hogy minden $0 < p \leq \frac{1}{2}$ $0 < \varepsilon \leq pq$ esetén fennáll a

$$(2.15) \quad C_*(\varepsilon, p) > C(\varepsilon, p)$$

reláció. Ennek érdekében képezzük először $C_*(\varepsilon, p)$ hatványsorát:

$$\begin{aligned} C_*(\varepsilon, p) &= (\varepsilon + p) \ln\left(1 + \frac{\varepsilon}{p}\right) + (\varepsilon - q) \ln\left(1 + \frac{\varepsilon}{q - \varepsilon}\right) = \\ &= (\varepsilon + p) \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \left(\frac{\varepsilon}{p}\right)^j - (q - \varepsilon) \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \left(\frac{\varepsilon}{q - \varepsilon}\right)^j. \end{aligned}$$

Innen ε hatványai szerinti rendezés után adódik

$$C_*(\varepsilon, p) = \sum_{j=2}^{\infty} (-1)^j \varepsilon^j a_j,$$

ahol

$$a_j = a_j(\varepsilon, p) = \frac{(q - \varepsilon)^{j-1} + (j - 1)p^{j-1}}{j(j - 1)p^{j-1}(q - \varepsilon)^{j-1}}.$$

Egyszerűen belátható, hogy $C_*(\varepsilon, p)$ fenti előállítása $0 < p \leq \frac{1}{2}$ $0 < \varepsilon \leq pq$ esetén *Leibniz-féle alternáló sor*. Tekintsük ugyanis tetszőleges j index esetén a sor két egymást követő eleme abszolút értékének hányadosát:

$$q_j = \frac{a_{j+1}\varepsilon^{j+1}}{a_j\varepsilon^j} = \frac{\varepsilon}{p} \cdot \frac{j-1}{j+1} \cdot \frac{1 + j\left(\frac{p}{q-\varepsilon}\right)^j}{1 + (j-1)\left(\frac{p}{q-\varepsilon}\right)^{j-1}}.$$

Mivel tetszőleges $u > v > 0$ és $w > 0$ esetén $\frac{u+w}{v+w} < \frac{u}{v}$, ezért a $\frac{p}{q-\varepsilon} > 1$ esetben

$$q_j < \frac{\varepsilon}{p} \cdot \frac{j-1}{j+1} \cdot \frac{j\left(\frac{p}{q-\varepsilon}\right)^j}{(j-1)\left(\frac{p}{q-\varepsilon}\right)^{j-1}} = \frac{\varepsilon}{q-\varepsilon} \cdot \frac{j}{j+1} < 1$$

(hiszen $0 < p \leq \frac{1}{2}$ mellett $\varepsilon \leq pq \leq \frac{q}{2}$).

A $\frac{p}{q-\varepsilon} \leq 1$ esetben viszont nyilvánvaló, hogy

$$q_j \leq \frac{\varepsilon}{p} \cdot \frac{j-1}{j+1} \cdot \frac{1+j\left(\frac{p}{q-\varepsilon}\right)^{j-1}}{1+(j-1)\left(\frac{p}{q-\varepsilon}\right)^{j-1}} < \frac{\varepsilon}{p} \cdot \frac{j-1}{j+1} \cdot \frac{j}{j-1} < 1,$$

tehát $\sum_{j=2}^{\infty} (-1)^j \varepsilon^j a_j$ valóban *Leibniz-féle alternáló sor*. Innen adódik az

$$(2.16) \quad \varepsilon^2 a_2 - \varepsilon^3 a_3 < C_*(\varepsilon, p) < \varepsilon^2 a_2$$

becslés.

Ezek után annak igazolására, hogy minden $0 < p \leq \frac{1}{2}$ és $0 < \varepsilon \leq pq$ esetén teljesül (2.15), vagyis (lásd (2.13)–(2.14))

$$(2.17) \quad n_*(p, \varepsilon, \delta) < n(p, \varepsilon, \delta),$$

(2.16) alapján elég azt megmutatni, hogy fennáll

$$(2.18) \quad \varepsilon^2 a_2 - \varepsilon^3 a_3 - C(\varepsilon, p) > 0.$$

Vezessük be az $u = \frac{\varepsilon}{pq}$ jelölést, nyilván $0 < u \leq 1$. Ekkor

$$\begin{aligned} \varepsilon^2 a_2 - \varepsilon^3 a_3 - C(\varepsilon, p) &= \frac{\varepsilon^2}{2p} \left[1 + \frac{p}{q-\varepsilon} - \frac{\varepsilon \left[1 + 2 \left(\frac{p}{q-\varepsilon} \right)^2 \right]}{3p} - \frac{1}{q \left(1 + \frac{\varepsilon}{2pq} \right)^2} \right] = \\ &= \frac{\varepsilon^2}{2pq} \left[\frac{1-upq}{1-up} - \frac{uq}{3} \cdot \frac{(q-upq)^2 + 2p^2}{q(1-up)^2} - \frac{1}{\left(1 + \frac{u}{2} \right)^2} \right] = \frac{\varepsilon^2}{2pq} \cdot \frac{1}{3(1-up)^2 \left(1 + \frac{u}{2} \right)^2} \cdot \\ &\cdot \left\{ 3(1-upq)(1-up) \left(1 + \frac{u}{2} \right)^2 - u \left(1 + \frac{u}{2} \right)^2 [(q-upq)^2 + 2p^2] - 3(1-up)^2 \right\}. \end{aligned}$$

Itt az első két tényező pozitív, a kapcsos zárójelben szereplő kifejezés pozitivitása pedig egyszerű számolással adódik. Ezért fennáll (2.18), vagyis (2.17) is, tehát a leírt módszerrel a *Csebishev-egyenlőtlenségnek* egy, a *Bernstein-féle élesítésnél* valamivel jobb élesítése adódik a vizsgált speciális esetben.

3. Példák

Az alábbiakban a *Csebishev-egyenlőtlenség* leírt élesítése és a *Bernstein-féle élesítés* összehasonlítására néhány számpéldát mutatunk be. Az új becslés relatív hatékonyságát az $\frac{n_*}{n}$ hányadossal mérjük, ahol $n_* = n_*(p, \varepsilon, \delta)$ és $n = n(p, \varepsilon, \delta)$ a

$$P \left(\left| \frac{\xi^{*(n)}}{n} - p \right| \geq \varepsilon \right) \leq \delta \quad (0 < \varepsilon \leq pq, \quad 0 < \delta < 1)$$

reláció biztosításához elégséges lépésszámok (vö. (2.13)–(2.14)). Mivel

$$\frac{n_*}{n} = \frac{C(\varepsilon, p)}{C_*(\varepsilon, p)}$$

értéke δ -tól nem függ, ezért csak a p és ε értékeket tüntetjük fel.

p	ε	n_*/n
0,01	0,001	0,937
	0,002	0,878
	0,005	0,735
	0,009	0,597
0,05	0,005	0,931
	0,010	0,868
	0,025	0,718
	0,045	0,575
0,10	0,01	0,924
	0,02	0,856
	0,05	0,695
	0,09	0,546
0,20	0,02	0,907
	0,04	0,827
	0,10	0,644
	0,18	0,517
0,50	0,01	0,980
	0,02	0,924
	0,05	0,825
	0,10	0,690
	0,25	0,425

A számpéldákból látható, hogy az $\frac{n_*}{n}$ hányados értéke p vagy ε növekedtével egyaránt csökken, tehát az új becslés relatív hatékonysága nő. (Ezen belül az ε -tól, pontosabban az $u = \frac{\varepsilon}{pq}$ hányadostól való függés lényegesebb.)

4. Zajjal torzított függvényértékek közelítő meghatározása

Tegyük fel, hogy a $h(\mathbf{x}, \mathbf{y})$ ($\mathbf{x} \in E^n, \mathbf{y} = \boldsymbol{\eta}(\omega) \in E^q$) mérhető függvénnyel képezett

$$(4.1) \quad H(\mathbf{x}) = M(h(\mathbf{x}, \boldsymbol{\eta}))$$

korlátos függvény közelítő értékeit (különböző \mathbf{x} argumentumok mellett) az $\boldsymbol{\eta}$ -val azonos eloszlású, teljesen független $\boldsymbol{\eta}_k$ valószínűségi változók \mathbf{y}_k , $k=1, 2, 3, \dots$ realizációi alapján kívánjuk kiszámítani. A nagy számok erős törvénye ([2], 330. o.) értelmében

$$P\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n h(\mathbf{x}, \boldsymbol{\eta}_k) = H(\mathbf{x})\right) = 1;$$

a korábbiakhoz hasonlóan a

$$(4.2) \quad P\left(\left|\frac{1}{n} \sum_{k=1}^n h(\mathbf{x}, \boldsymbol{\eta}_k) - H(\mathbf{x})\right| \geq \varepsilon\right) \leq \delta \quad (\varepsilon > 0, \quad 0 < \delta < 1)$$

reláció kielégítéséhez elegendő n érték becslését keressük. Az n értéket általában a centrális határeloszlás-tétel alapján szokták (közelítőleg) meghatározni: itt a 2. szakaszban leírt módszer kiterjesztése alapján adunk n -re vonatkozó becslést.

Vezessük be a

$$(4.3) \quad \begin{aligned} \xi_k^*(\mathbf{x}) &= h(\mathbf{x}, \boldsymbol{\eta}_k) \quad G_x^*(z) = P(\xi_k^*(\mathbf{x}) < z), \quad k = 1, 2, 3, \dots \\ \xi_k(\mathbf{x}) &= \xi_k^*(\mathbf{x}) - M(\xi_k^*(\mathbf{x})) = \xi_k^*(\mathbf{x}) - H(\mathbf{x}), \quad G_x(z) = P(\xi_k(\mathbf{x}) < z), \\ &k = 1, 2, 3, \dots \end{aligned}$$

$$\xi^{*(n)}(\mathbf{x}) = \sum_{k=1}^n \xi_k^*(\mathbf{x}), \quad \xi^{(n)}(\mathbf{x}) = \sum_{k=1}^n \xi_k(\mathbf{x}), \quad G_x^{(n)}(z) = P(\xi^{(n)}(\mathbf{x}) < z), \quad n = 1, 2, 3, \dots$$

jelöléseket; a továbbiakban — ahol ez nem okoz félreértést — a (4.3)-ban definiált $\xi_k^*(\mathbf{x})$, $\xi_k(\mathbf{x})$, $\xi^{*(n)}(\mathbf{x})$, $\xi^{(n)}(\mathbf{x})$ valószínűségi változók és a $G_x^*(z)$, $G_x(z)$, $G_x^{(n)}(z)$ eloszlásfüggvények \mathbf{x} paraméterének jelölését elhagyjuk.

4.1. LEMMA. Tegyük fel, hogy $\xi^{(n)}$ eloszlása minden n -re szimmetrikus, és $0 < v \leq V$ esetén létezik $M(e^{v\xi_k^*})$. Ekkor tetszőleges $0 < v \leq V$ és $\varepsilon > 0$ mellett fennáll a

$$(4.4) \quad P\left(\left|\frac{\xi^{*(n)}}{n} - H(\mathbf{x})\right| \geq \varepsilon\right) \leq 2e^{-n[v(\varepsilon + H(\mathbf{x})) - \ln M(e^{v\xi_k^*})]}$$

reláció.

Bizonyítás. $\xi^{(n)}$ eloszlásának szimmetriája miatt felírható:

$$(4.5) \quad \begin{aligned} P\left(\left|\frac{\xi^{*(n)}}{n} - H(\mathbf{x})\right| \geq \varepsilon\right) &= P(|\xi^{(n)}| \geq n\varepsilon) \leq \\ &\leq P(\xi^{(n)} \geq n\varepsilon) + P(-\xi^{(n)} \geq n\varepsilon) = 2P(\xi^{(n)} \geq n\varepsilon). \end{aligned}$$

Ezért az

$$\varepsilon = \frac{t + \ln M(e^{v\xi_k^*})}{nv} > 0 \quad (t > 0)$$

jelöléssel, a 2.3. segédteétel igazolásához analóg módon adódik

$$(4.6) \quad \begin{aligned} P(\xi^{(n)} \geq n\varepsilon) &= P\left(\xi^{(n)} \geq \frac{t + \ln M(e^{v\xi_k^*})}{v}\right) \leq e^{-t} = \\ &= e^{-nve + \ln M(e^{v\xi_k^*})} = e^{-nve + n \ln M(e^{v\xi_k^*})} = \\ &= e^{-n[v\varepsilon - \ln M(e^{v(\xi_k^* - H(\mathbf{x}))})]} = e^{-n[v(\varepsilon + H(\mathbf{x})) - \ln M(e^{v\xi_k^*})]}, \end{aligned}$$

tehát (4.5) és (4.6) alapján fennáll a (4.4) reláció.

4.1. Megjegyzés. A $\xi^{(n)}$ eloszlásának szimmetriájára vonatkozó feltétel teljesül, ha a $\xi_k^*(\mathbf{x})$, $k=1, 2, 3, \dots$ valószínűségi változók eloszlása ($H(\mathbf{x})$ várható értékükre vonatkozóan) szimmetrikus. Nyilvánvaló, hogy ehelyett elég megkövetelni pl. azt,

hogy minden z -re teljesüljön

$$P(\xi^{(n)} \cong z) \cong P(-\xi^{(n)} \cong z), \quad n = 1, 2, 3, \dots$$

A továbbiakban a (4.4) típusú becslés élesítésére törekszünk. Legyen (rögzített $\varepsilon > 0$ és $H(\mathbf{x})$ értékek mellett)

$$F(v) = v(\varepsilon + H(\mathbf{x})) - \ln M(e^{v\xi_k^*}) = \frac{t}{n} > 0.$$

Ha az $F(v)$ függvénynek létezik maximumhelye a $(0, V]$ intervallumon, akkor az a

$$\frac{\partial F(v)}{\partial v} = \varepsilon + H(\mathbf{x}) - \frac{\partial}{\partial v} M(e^{v\xi_k^*}) \cdot \frac{1}{M(e^{v\xi_k^*})} = 0$$

egyenlet alapján határozható meg. Feltéve, hogy

$$M(e^{v\xi_k^*}) = \int_{-\infty}^{\infty} e^{vz} dG^*(z) \quad (0 < v \leq V)$$

v szerinti differenciálása az integrálás műveletén belül elvégezhető, adódik

$$\frac{\partial}{\partial v} M(e^{v\xi_k^*}) = M(e^{v\xi_k^*} \cdot \xi_k^*).$$

Az elmondottak alapján érvényes a következő állítás:

4.2. LEMMA.

Ha a

$$K(v) = \frac{M(e^{v\xi_k^*} \cdot \xi_k^*)}{M(e^{v\xi_k^*})} \quad (0 < v \leq V)$$

függvény $v > 0$ esetén monoton növekvő, és létezik olyan $0 < v_* \leq V$ érték, amellyel fennáll a

$$(4.7) \quad K(v_*) = \varepsilon + H(\mathbf{x})$$

reláció, akkor v_* az $F(v)$ függvény maximumhelye a $(0, V]$ intervallumon.

A (4.2), (4.4) és (4.7) relációk egybevetésével adódik az alábbi

4.1. TÉTEL. Tegyük fel, hogy fennállnak a 4.1. és 4.2. segédtelemek feltételei. Ekkor a (4.1)-ben definiált $H(\mathbf{x})$ függvényértéknek a (4.2) relációt kielégítő pontoságú becsléséhez elegendő a $\xi_k^*(\mathbf{x}) = h(\mathbf{x}, \eta_k)$, $k = 1, 2, 3, \dots$ valószínűségi változók

$$n \cong n_*(H(\mathbf{x}), \varepsilon, \delta) = \ln \frac{2}{\delta} \cdot \frac{1}{v_*(\varepsilon + H(\mathbf{x})) - \ln M(e^{v_*\xi_k^*})}$$

számú független realizációja; itt v_* a (4.7) egyenlet megoldását jelöli.

4.2. Megjegyzés. A 4.2. lemmában szereplő $K(v)$ függvény monotonitásának szükséges és elégséges feltétele a $\frac{\partial K(v)}{\partial v} \cong 0$ ($v > 0$) reláció alapján az

$$(4.8) \quad \int_{-\infty}^{\infty} e^{vz} dG^*(z) \cdot \int_{-\infty}^{\infty} e^{vz} z^2 dG^*(z) \cong \left[\int_{-\infty}^{\infty} e^{vz} z dG^*(z) \right]^2 \quad (v > 0)$$

egyenlőtlenség fennállása (feltéve, hogy a szereplő integrálok léteznek).

Vizsgáljuk először a (4.8) feltételt véges értékű készletű ξ_k^* esetén, vagyis legyen

$$P(\xi_k^* = z_i) = p_i, \quad i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1.$$

Ekkor (4.8) a

$$\left(\sum_{i=1}^m p_i e^{v z_i} \right) \left(\sum_{i=1}^m p_i e^{v z_i} z_i^2 \right) \geq \left(\sum_{i=1}^m p_i e^{v z_i} z_i \right)^2$$

alakban írható fel: ez viszont mindig igaz, hiszen

$$\begin{aligned} & \sum_{i=1}^m \sum_{j=1}^m p_i p_j e^{v z_i} e^{v z_j} z_j^2 - \sum_{i=1}^m \sum_{j=1}^m p_i p_j e^{v z_i} z_i e^{v z_j} z_j = \\ &= \sum_{i=1}^m \sum_{j=1}^m p_i p_j e^{v(z_i+z_j)} (z_j^2 - z_i z_j) = \sum_{1 \leq i < j \leq m} p_i p_j e^{v(z_i+z_j)} (z_i^2 + z_j^2 - 2z_i z_j) \geq 0. \end{aligned}$$

Mivel tetszőleges $G^*(z)$ eloszlásfüggvény előállítható lépcsős függvények limeszeként, így — a szereplő integrálok létezése esetén — mindig igaz a (4.8) reláció. Ez azt jelenti, hogy a 4.1. és 4.2. lemmák további feltételei mellett a leírt módszer (legalábbis elvben) alkalmazható a (4.2) reláció fennállását biztosító n érték meghatározására.

4.3. *Megjegyzés.* Az elmondottakból látható, hogy n értékére vonatkozóan akkor is tudunk becslést adni, ha v_* értéke ugyan nem ismert, viszont (a vizsgált x és pl. alkalmas $v > 0$ esetén) meghatározható az $F(v)$ függvényérték egy (pozitív) alsó becslése. A szóban forgó becslés értékét $C(x, \varepsilon, v)$ -vel jelölve, (4.2) és (4.4) összevetése alapján látható, hogy

$$n \geq \ln \frac{2}{\delta} \frac{1}{C(x, \varepsilon, v)}$$

elégseges ahhoz, hogy fennálljon a (4.2) reláció.

Illusztrációképpen tekintsük azt az esetet, amikor $0 \leq v \leq V_1(x)$ mellett

$$(4.9) \quad M(e^{v \xi_k^*(x)}) \leq e^{\frac{1}{2} B(x) v^2},$$

ahol $V_1(x)$ és $B(x)$ az x paramétertől függő pozitív állandók. Tegyük fel még, hogy $0 < \varepsilon \leq B(x) \cdot V_1(x)$. Ekkor, mivel

$$\begin{aligned} F(v) &= v(\varepsilon + H(x)) - \ln M(e^{v \xi_k^*(x)}) v = \varepsilon - \ln M(e^{v \xi_k^*(x)}) \geq \\ &\geq v\varepsilon - \frac{1}{2} B(x) v^2 = \hat{F}(v), \quad 0 \leq v \leq V_1(x), \end{aligned}$$

ezért az $\hat{F}(v)$ kifejezést maximalizáló $\hat{v} = \frac{\varepsilon}{B(x)} > 0$ választással látható, hogy

$$(4.10) \quad n \geq n_*(B(x), \varepsilon, \delta) = \ln \frac{2}{\delta} \cdot \frac{1}{\hat{F}(\hat{v})} = \ln \frac{2}{\delta} \cdot \frac{2B(x)}{\varepsilon^2}$$

$$(0 < \varepsilon \leq B(x) \cdot V_1(x), \quad 0 < \delta < 1)$$

esetén teljesül a (4.2) reláció. Nem nehéz belátni, hogy a (4.9) feltétel ekvivalens az

$$M(e^{v \xi_k^*(x)}) < \infty, \quad 0 \leq v \leq V_2, \quad V_2 = V_2(x) > 0$$

feltétellel; ekkor $B(x)$ értéként tetszőleges, az $M(\xi_k^2(x))$ értékénél nagyobb szám választható. Ez az eredmény összhangban van az 1. és 2. szakaszban felsorolt, illetve bizonyított (az ismeretlen valószínűségek becsléséhez elégséges n értékekre vonatkozó) eredményekkel, és azokat bizonyos értelemben általánosítja. Ha ugyanis speciálisan ξ_k^* , $k=1, 2, 3, \dots$ egy $0 < p < 1$ ($q=1-p$) valószínűségű esemény indikátorai, akkor a $B(x)=B_i$, $i=1, 2, 3, 4$ definíciókkal

$$B_1 = \frac{pq}{2\delta \ln \frac{2}{\delta}} \quad \left(0 < \delta < 1, 0 < 2\delta \ln \frac{2}{\delta} < 1\right)$$

$$B_2 = \left(1 + \frac{\varepsilon}{2pq}\right)^2 pq \quad (0 < \varepsilon \leq pq)$$

$$B_3 = \frac{k^2(\delta)}{2\ln \frac{2}{\delta}} pq \quad \left(0 < \delta < 1, k^2(\delta) > 2\ln \frac{2}{\delta}\right)$$

$$B_4 = \frac{\varepsilon^2}{2\ln \left[\left(\frac{\varepsilon+p}{p}\right)^{\varepsilon+p} \left(\frac{q}{q-\varepsilon}\right)^{\varepsilon-q}\right]} \quad \left(0 < \varepsilon \leq pq, p \leq \frac{1}{2}\right)$$

esetén (4.10)-ből rendre a *Csebisev-egyenlőtlenséget*, annak *Bernstein-féle élesítését*, a centrális határeloszlás-tétel alapján adódó közelítést, illetve a (2.13) relációt nyerjük.

IRODALOM

- [1] DEÁK, I., „Monte Carlo-módszerek a több dimenziós térben elhelyezkedő halmazok valószínűségének meghatározására normális eloszlás esetén”, *Alkalmazott Matematikai Lapok* 4 (1978) 35–94.
- [2] RÉNYI, A., *Valószínűségszámítás* (Tankönyvkiadó, Budapest, 1968).
- [3] БЕРНШТЕЙН, С. Н., Теория вероятностей (4-е издание, Гостехиздат, Москва, 1946).
- [4] БУСЛЕНКО, Н. П., Моделирование сложных систем (2-е издание, Наука, Москва, 1978).
- [5] Ермаков, С. М., Метод Монте-Карло и смежные вопросы (Наука, Москва, 1975).
- [6] ШИРЯЕВ, А. Н., Вероятность (Наука, Москва, 1980).

(Beérkezett: 1982. július 23.)

PINTÉR JÁNOS

ELTE SZÁMÍTÓKÖZPONT, OPERÁCIÓKUTATÁSI OSZTÁLY
1093 BUDAPEST, DIMITROV TÉR 8.

AN IMPROVED CHEBYSHEV-INEQUALITY FOR FUNCTION VALUE ESTIMATES BY MONTE CARLO TECHNIQUES

J. PINTÉR

An improved *Chebyshev-type inequality* is proved and compared to the *Chebyshev—Bernstein inequality*, in the case of relative frequency estimates of unknown probabilities. The degree of improvement is analyzed and illustrated by numerical examples. An extension of the described technique for evaluating noise-perturbed function values is also presented.

A LEGERŐSEBB PRÓBÁK TÉMAKÖRÉNEK MATEMATIKAI MEGKÖZELÍTÉSE

BÍRÓ MIKLÓS

Budapest

A *Neyman—Pearson lemma* a matematikai statisztika egyik alapvető eredménye. A dolgozat e lemmára, valamint a monoton valószínűséghányadosú eloszlásösszességek esetén ismert állításra ad matematikai programozási magyarázatot. Felvetődik a nem véletlenített próbák konstruálásának lehetősége 0—1 diszkrét programozás felhasználásával.

1. Bevezetés

Az [1] dolgozatban leírt gondolatok kiindulópontjául PRÉKOPA ANDRÁS [8] cikke és szóbeli ösztönzése szolgált.

Az [1] dolgozat bevezetésében az [5] könyv egy részletének idézete bizonyítja azt, hogy a legjobb statisztikai próbák konstruálásakor felvetődött a diszkrét programozási probléma. Ezt azonban folytonos relaxációval, vagy a statisztika nyelvén véletlenítéssel oldották meg. Az így kapott feladat véges nullhipotézis és egyszerű ellenhipotézis esetén egy lineáris programozási probléma. A következő szakasz célja e feladat bevezetése. Az előbbi feltételek mellett a *Neyman—Pearson lemma* egy ismert lineáris programozási tételnek felel meg.

Az utolsó szakaszban a *Neyman—Pearson-lemma* monoton valószínűséghányadosú eloszlásösszességek esetén ismert alakját világítjuk meg lineáris programozási szempontból. Itt az ellenhipotézis nem feltétlenül egyszerű, ami továbbra is — véges hipotézisek esetén — egy több célfüggvényű lineáris programozási feladathoz vezet. A monotonitási feltevés miatt azonban e feladat megoldása pontosan olyan alakú, mint az egy célfüggvényű és egy feltételes ún. hátizsák feladat megoldása.

A teljesség kedvéért definiáljuk a továbbiakban felhasználandó fogalmakat ([10]).

Legyen egy kísérlet X kimenetele az M mintatér eleme.

Tekintsünk egy véges sok elemből álló eloszlásösszességet, az (M, β) mérhető téren. X alapján döntenünk kell a következő hipotézisek között

H_0 : Az X minta a

$$\{P_{\mathfrak{g}}: \mathfrak{g} \in \{\mathfrak{g}_1, \dots, \mathfrak{g}_m\}\}$$

eloszlásösszesség valamelyik eleméből származik.

H_1 : Az X minta a

$$P_{\mathfrak{g}_{m+1}}$$

eloszlásból származik.

(Pillanatnyilag csak egyszerű ellenhipotéziseket vizsgálunk. A nullhipotézis lehet összetett is, de pillanatnyilag csak véges sok elemből álló eloszlásösszességre vonatkozhat.)

Egy próba konstrukciója az M mintatér egy ún. kritikus tartományának, K -nak a megadását jelenti. A nullhipotézist pontosan akkor utasítjuk el, ha az X minta a K tartományba esik. A K tartomány φ indikátorfüggvényét nevezzük próbafüggvénynek:

$$\varphi(X) = \begin{cases} 1, & \text{ha } X \in K, \\ 0, & \text{ha } X \notin K. \end{cases}$$

Egy próba alkalmazása során kétféle hibát követhetünk el. Elsőfajú hibának nevezik azt, amikor a valóságban fennáll a H_0 hipotézis, a kísérlet eredménye mégis a kritikus tartományba esik, ezért a H_0 hipotézist elutasítjuk. Másodfajú hibát követünk el, ha a H_0 hipotézis a valóságban nem áll fenn, mégis elfogadjuk, mert a kísérlet eredménye nem esett a kritikus tartományba. Erőfüggvénynek nevezzük a másodfajú hiba kiegészítő eseményének a valószínűségét. Ez tehát a kritikus tartományba esés valószínűsége abban az esetben, amikor a H_0 hipotézis nem áll fenn.

Egy K kritikus tartománnyal definiált próbát α terjedelmű egyenletesen legerősebb próbának nevezünk, ha amellet a feltétel mellett, hogy az elsőfajú hiba valószínűsége nem haladja meg az α -t, a próba ereje maximális.

2. A Neyman—Pearson-lemma

Az α terjedelmű egyenletesen legerősebb próba meghatározása a következő feladatnak felel meg:

$$\max P_{\vartheta_{m+1}}(K)$$

$$P_{\vartheta_i}(K) \leq \alpha, \quad i = 1, \dots, m,$$

vagy a φ indikátorfüggvény várható értékét tekintve:

$$(2.1) \quad \max \int_M \varphi dP_{\vartheta_{m+1}}$$

$$\int_M \varphi dP_{\vartheta_i} \leq \alpha, \quad i = 1, \dots, m$$

$$\varphi(X) \in \{0, 1\} \quad \forall X \in M.$$

Ha az M mintatér, azaz a P eloszlások végesek ($M = \{x_1, \dots, x_n\}$), akkor ez a feladat a következő alakot ölti:

$$\max \sum_{j=1}^n \varphi(x_j) P_{\vartheta_{m+1}}(X = x_j)$$

$$(2.2) \quad \sum_{j=1}^n \varphi(x_j) P_{\vartheta_i}(X = x_j) \leq \alpha,$$

$$i = 1, \dots, m,$$

$$\varphi(x_j) \in \{0, 1\},$$

$$j = 1, \dots, n.$$

Mint láthatjuk, ez egy 0–1-es lineáris programozási feladat.

A statisztikai elméletben az ekkor szokásos lépés a próbafüggvény véletlenítése, ami matematikai programozási terminológiával a diszkrét programozási feladat folytonos relaxálását jelenti a $\varphi(x_j) \in \{0, 1\}$ feltétel helyett a $0 \leq \varphi(x_j) \leq 1$ feltétel elfogadásával. A véletlenített próbafüggvénynek egy X minta esetén felvett $\varphi(X) = p$ értéke azt jelenti, hogy az X mintát p valószínűséggel tekintjük a kritikus tartomány elemének.

Egyébként is véletlen jelenségekről lévén szó, a próba véletlenítésének elfogadása indokolt. Azonban a véletlenítés elfogadásának olyan technikai okai is voltak [5], amelyek ma már bizonyos esetekben megszűntek. Itt a diszkrét programozási feladat megoldási módszereire gondolok. Egyszerű nullhipotézis esetén bináris hátizsák feladattal állunk szemben, tehát alkalmazhatók az [1] dolgozatban leírt eljárások. Lehetőség van tehát legjobb nem véletlenített próbák konstruálására is. Ezeknek a vizsgálata és gyakorlati összehasonlítása a legjobb véletlenített próbákkal nem tartozik e dolgozat célkitűzései közé.

Most vizsgáljuk meg, hogyan következik a *Neyman—Pearson-lemma* egy általánosítása a matematikai programozás elméletéből. A történeti hűség érdekében azonban DANTZIG [3] könyvére hivatkozva, inkább a következő tételt kell a későbbi lineáris programozási kutatások egyik alapjának tekinteni.

2.1. TÉTEL [5, 11]. A fenti jelölésekkel, véges M mintater esetén az α terjedelmű φ próbafüggvény akkor és csak akkor határozza meg a H_0 hipotézisnek H_1 -gyel szembeni legerősebb próbáját, ha létezik egy olyan $q \in R^m$ vektor, amelyre

1.

$$q_i \left(\alpha - \sum_{j=1}^n \varphi(x_j) P_{g_i}(X = x_j) \right) = 0 \quad i = 1, \dots, m,$$

2. $\varphi(x_j) = 1$, ha

$$P_{g_{m+1}}(X = x_j) > \sum_{i=1}^m q_i P_{g_i}(X = x_j)$$

3. $\varphi(x_j) = 0$, ha

$$P_{g_{m+1}}(X = x_j) < \sum_{i=1}^m q_i P_{g_i}(X = x_j)$$

A 2.1. tétel a következő lineáris programozási tétel speciális esete.

2.2. TÉTEL (*felső korlátos LP optimális megoldása*). Tekintsük a

(2.3)

$$\max c'x$$

$$Ax \leq b$$

$$0 \leq x \leq U$$

feladat egy x megengedett megoldását. Jelöljük a_j -vel az A mátrix j -edik oszlopát.

Az x vektor akkor és csak akkor optimális megoldása a (2.3) feladatnak, ha

létezik egy olyan $\mathbf{q} \neq \mathbf{0}$ vektor, amelyre

1. $(\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{q} = 0$,
2. $c_j - \mathbf{q}'\mathbf{a}_j > 0$ esetén $x_j = U_j$,
3. $c_j - \mathbf{q}'\mathbf{a}_j < 0$ esetén $x_j = 0$.

A *Neyman—Pearson-lemma* eredeti alakja [5, 6, 7, 8, 10, 11], amelyben mindkét hipotézis egyszerű, a folytonos hátizsák feladat optimális megoldásának felel meg ([8]). Ez utóbbi is nyilvánvalóan a 2.2. tétel következménye ([1]).

3. Monoton valószínűséghányadosú eloszlásösszesség esete

Felmerül a kérdés, hogy vajon nem egyszerű, esetleg végtelen hipotézisek esetén felhasználható-e a 2.1. tétel.

A matematikai statisztikában ismert válasz, hogy egyparaméteres monoton valószínűséghányadosú eloszlásösszesség esetén felhasználható ([10, 11]).

3.1. DEFINÍCIÓ. Legyen T egy valószínű statisztika. A

$$\{P_{\vartheta} : \vartheta \in \Theta \subset R\}$$

eloszlásösszességet (ahol P_{ϑ} vagy sűrűségfüggvény, vagy diszkrét eloszlás esetén valószínűség) $T(x)$ -ben monoton valószínűséghányadosúnak nevezzük, ha $\vartheta_1 < \vartheta_2$ esetén létezik egy $f_{\vartheta_1, \vartheta_2}(t)$ monoton nem csökkenő függvény, amelyre

$$\frac{P_{\vartheta_1}(x)}{P_{\vartheta_2}(x)} = f_{\vartheta_1, \vartheta_2}(T(x)).$$

Véges eloszlások esetén ez a definíció azt jelenti, hogy $\vartheta_{i_1} < \vartheta_{i_2}$ esetén a

$$\frac{P_{\vartheta_{i_1}}(X = x_j)}{P_{\vartheta_{i_2}}(X = x_j)}$$

hányadosok csökkenő sorrendje ϑ_{i_1} és ϑ_{i_2} konkrét értékétől független.

A 3.1. definíció első látásra túl szigorúnak tűnhet, azonban egy nagyon gyakori esetben teljesül az ott leírt feltétel. Ezaz egyparaméteres exponenciális eloszlás-családok esete, amelyeknek sűrűségfüggvénye a következő alakú:

$$P_{\vartheta}(x) = C(\vartheta)e^{D(\vartheta)T(x)}h(x).$$

Ide tartozik például a $B(n_1, 0)$ binomiális eloszlások családja adott n_1 esetén, az adott szórású normális eloszlások családja és az exponenciális eloszlások.

Most tekintsünk egy véges eloszlásokból álló

$$\{P_{\vartheta} : \vartheta \in \Theta \subset R\}$$

eloszlásösszességet. Egy kísérlet X kimenetele alapján döntenünk kell a következő összetett hipotézisek között:

H_0 : az X minta a

$$\{P_\vartheta: \vartheta \leq \vartheta_0\}$$

eloszlásösszesség valamelyik eleméből származik,

H_1 : az X minta a

$$\{P_\vartheta: \vartheta > \vartheta_0\}$$

eloszlásösszesség valamelyik eleméből származik.

Ha a Θ paraméterhalmaz véges, akkor az α terjedelmű egyenletesen legerősebb próba meghatározása a következő több célfüggvényű lineáris programozási feladat megoldásának felel meg:

$$\max \sum_{j=1}^n \varphi(x_j) P_\vartheta(X = x_j), \quad \forall \vartheta > \vartheta_0$$

$$(3.1) \quad \sum_{j=1}^n \varphi(x_j) P_\vartheta(X = x_j) \leq \alpha, \quad \forall \vartheta \leq \vartheta_0$$

$$0 \leq \varphi(x_j) \leq 1, \quad j = 1, \dots, n.$$

Most tekintsük a (3.1) feladatot monoton valószínűséghányadosú eloszlásösszesség esetén. A tárgyalás egyszerűsítése érdekében térjünk át a lineáris programozási jelölésrendszerre. Így a feladat a következő alakú lesz:

$$\max Cx$$

$$(3.2) \quad Ax \leq b$$

$$0 \leq x \leq 1,$$

ahol a c_{kj}/a_{ij} számok csökkenő sorrendje k és i konkrét értékétől független. Ez azt jelenti, hogy ha a változókat a C és A mátrix egy tetszőleges k , illetve i sorát kiválasztva a hányadosok csökkenő sorrendje szerint rendeztük, akkor ez a sorrend minden más célfüggvény-feltétel párja is fennáll.

3.1. TÉTEL. Tegyük fel, hogy a (3.2) feladat változóit a fentiek szerint rendeztük. Ekkor a (3.2) feladat x^* optimális megoldása :

$$x_j^* = \begin{cases} 1, & j = 1, \dots, p-1, \\ \frac{1}{a_{sp}} \left(b_s - \sum_{k=0}^{p-1} a_{sk} \right), & j = p, \\ 0, & j = p+1, \dots, n, \end{cases}$$

ahol

$$p = \min \left\{ k: \exists i: \sum_{j=1}^k a_{ij} > b_i \right\},$$

és s olyan, hogy

$$s \in \left\{ i: \sum_{j=1}^p a_{ij} > b_i \right\},$$

valamint

$$\frac{1}{a_{sp}} \left(b_s - \sum_{j=1}^{p-1} a_{sj} \right) \leq \frac{1}{a_{ip}} \left(b_i - \sum_{j=1}^{p-1} a_{ij} \right)$$

minden i indexre az

$$\left\{ i: \sum_{j=1}^p a_{ij} > b_i \right\}$$

halmazból.

Bizonyítás. Az \mathbf{x}^* vektor nyilván megengedett megoldása a (3.2) feladatnak, viszont minden k -ra optimális megoldása a következő hátizsák feladatnak:

$$\max \sum_{j=1}^n c_{kj} x_j$$

$$\sum_{j=1}^n a_{sj} x_j \leq b_s$$

$$0 \leq x_j \leq 1, \quad j = 1, \dots, n.$$

Mivel e feladat feltétele a (3.2) feladat feltételei között szerepel, ezért \mathbf{x}^* annak is optimális megoldása.

Végül vegyük észre, hogy a 3.1. tételben és annak bizonyításában nem játszik szerepet az \mathbf{A} mátrix sorainak száma. Így az akár végtelen, sőt nem megszámlálható is lehet. Ez utóbbi eset a nem véges sok elemből álló egyparaméteres eloszlásösszesítésnek felel meg.

Vegyük észre továbbá, hogy bár a tételeket véges M mintatér, azaz véges eloszlások esetére mondtuk ki, ezek dominált eloszlásösszesések esetén általánosíthatók. Ekkor a szummákat a domináló mérték szerinti integrálok, a valószínűségeket pedig sűrűségfüggvények helyettesítik ([2, 4, 5, 11]).

Végül szeretnék köszönetet mondani PRÉKOPA ANDRÁSNAK, hogy felhívta figyelmemet a hátizsák probléma megoldása és a Neyman—Pearson-lemma közötti kapcsolatra.

IRODALOM

- [1] BÍRÓ, M., „A bináris hátizsák feladat”. Megjelenés alatt az Alk. Mat. Lapokban.
- [2] DANTZIG, G. B. and WALD, A., “On the fundamental lemma of Neyman and Pearson”, *Annals of Mathematical Statistics* **22** (1951).
- [3] DANTZIG, G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, New Jersey, 1963).
- [4] FRANCIS, R. L. and WRIGHT, G. P., “Some duality relationships for the generalized Neyman—Pearson problem”, *Journal of Optimization Theory and Applications* **4** (1969) 394—412.
- [5] LEHMANN, E., *Testing Statistical Hypotheses* (John Wiley and Sons, New York, 1959).
- [6] NEYMAN, J. and PEARSON, E. S., “On the problem of the most efficient tests of statistical hypotheses”, *Philosophical Transactions of the Royal Society of London, Series A*, **231** (1933).

- [7] NEYMAN, J and PEARSON, E. S., "Contributions to the theory of testing statistical hypotheses", *Stat. Res. Memoirs* 1 (1936) 1—37.
- [8] PRÉKOPA, A., "The use of stochastic programming for the solution of some problems in probability and statistics", in: *Extremal Methods and Systems Analysis, Proceedings of an International Conference in honour of A. Charnes' sixtieth birthday*, Austin, Texas 1977, *Lecture Notes in Economics and Mathematical Systems*, 174, (Springer Verlag, 1980) 522—536.
- [9] PRÉKOPA, A., „A sztochasztikus programozás és a statisztika kapcsolata”, megjelenés alatt.
- [10] VINCZE, I., *Matematikai statisztika ipari alkalmazásokkal* (Műszaki Könyvkiadó, Budapest, 1968).
- [11] WITTING, H., *Mathematische Statistik* (B. G. Teubner, Stuttgart, 1966).

(Beérkezett: 1982. augusztus 12.)

BÍRÓ MIKLÓS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1502 BUDAPEST XI., KENDE U. 13—17.

ON A MATHEMATICAL PROGRAMMING APPROACH TO THE TOPIC OF THE MOST EFFICIENT TESTS

M. BÍRÓ

The *Lemma of Neyman and Pearson* is one of the basic results in the theory of mathematical statistics. This paper gives a mathematical programming explanation of this lemma and of the known theorem for the case of distribution families with monoton probability ratio. The possibility of the construction of non-randomized tests by 0—1 integer programming is raised.

A BINÁRIS HÁTIZSÁK FELADAT

BÍRÓ MIKLÓS

Budapest

A dolgozat áttekintést nyújt a bináris hátizsák feladat megoldásának alapvető módszereiről. Ezenbelül tartalmazza a szerzőnek a témában elért főbb eredményeit. Egyes tételek az itt szükségessé váló általánosabb, de máshol még használható és használt formában vannak megfogalmazva és bizonyítva. Ott, ahol a hátizsák feladat egyszerűbb bizonyítást is megenged, ezt is leírjuk. Néhány alkalmazási lehetőség után számítógépes tapasztalatokról is beszámolunk.

1. Bevezetés

A bináris hátizsák feladat a [28] könyv szerinti jelölésekkel a következőképpen írható fel:

$$(1.1) \quad \begin{aligned} &\max c'x \\ &a'x \leq b \\ &x \in \{0, 1\}^n \end{aligned}$$

vagy másképp

$$0 \leq x \leq 1 \quad \text{és} \quad x \text{ egész.}$$

A diszkrét feladat folytonos megfelelőjének nevezzük a következőt:

$$(1.2) \quad \begin{aligned} &\max c'x \\ &a'x \leq b \\ &0 \leq x \leq 1. \end{aligned}$$

Az (1.2) feladat egy lineáris programozási probléma, amely azonban csak egy feltételt tartalmaz (az $x \leq 1$ feltételeket implicit módon vesszük figyelembe). Ez a tulajdonság olyan megoldó algoritmus készítését teszi lehetővé, amelynek hatékonysága lényegesen jobb a több feltételes feladatok megoldására ismert simplex módszerénél. A gondolat érdekes módon két egymástól független területen vetődött fel, melyek közül az egyik természetesen a lineáris programozás elmélete.

A folytonos hátizsák feladat ma már közismert megoldási módszerét DANTZIG 1957-ben megjelent cikkében [6] írta le. Ez a megoldás azonban a matematika egy másik területén, a statisztikai hipotézisek vizsgálatának elméletében már 1933 óta ismert volt. Ekkor jelent meg NEYMAN és PEARSON munkája [27], amely ezen elmélet egyik alapvető eredményét tartalmazza. Az ismert lemma diszkrét eloszlások esetén pontosan a folytonos hátizsák feladat optimális megoldását adja meg. A lemma általánosítása összetett nullhipotézisre diszkrét eloszlások esetén szintén levezethető a lineáris programozás elméletéből.

Az (1.1) feladat megoldása lényegesen nehezebb, mint a folytonos megfelelőjéé. (1.1) az NP-teljes feladatok családjába tartozik.

A diszkrét probléma tulajdonképpen már a *Neyman—Pearson-probléma* vizsgálata során is felmerül. Itt a nehézséget csak a véletlenített próbák bevezetésével sikerült elkerülni, ami lényegében a diszkrét optimum helyett a folytonos elfogadását jelenti. LEHMANN 1959-ben megjelent *Testing Statistical Hypothesis* című könyvében [24] a 64. oldalon jól írja le a diszkrét feladat megoldása során fellépő legfőbb nehézséget.

„It may happen that when a certain point is included, the value α has not yet been reached but that it would be exceeded if the next point were also included. The exact value α can then either not be achieved at all, or it can be attained only by passing over the next desirable point and in its place taking one further down the list. The difficulty can be overcome by permitting randomization.”

Itt α az (1.1) feladat b jobb oldalának felel meg. LEHMANN feltehetően nem ismerte az éppen abban az időben az operációkutatás területén felszínre kerülő diszkrét programozási problémát [6].

A dolgozatban olyan módszerekkel foglalkozunk, amelyek segítségével hatékonyan lehet megoldani bináris hátizsák feladatokat.

A dolgozat megírásának egyik alapelve az volt, hogy az csak magyar nyelvű irodalom alapján is követhető legyen, ezért ahol lehetséges, ott magyar nyelvű irodalomra hivatkozunk.

A második szakaszban a folytonos feladat megoldására adunk egy mélyebb eszközöket felhasználó, valamint egy elemi megközelítést. Az azt követő szakaszban a bináris feladat megoldásához használt módszerekről és eszközökről lesz szó. A negyedik szakasz egy újabban felmerült algoritmikus probléma kiküszöbölésének lehetőségeit tárgyalja. Az ötödik szakaszban az alkalmazási lehetőségeket vázoljuk. Végül számítógépes tapasztalatokról is beszámolunk.

A jelölések általában megegyeznek PRÉKOPA ANDRÁS [28] könyvének jelöléseivel. Ezeken felül még a következőket használjuk általánosan:

+ diszjunkt halmazok uniója,

: = új jelölés bevezetése, vagy értékadás programozási értelemben.

2. A folytonos feladat

Annak ellenére, hogy fő célunk a diszkrét feladat megoldása, szükségünk van a folytonos megfelelőjének vizsgálatára is. Ennek oka, hogy a következő szakaszokban tárgyalandó módszerek vagy nagyszámú folytonos feladat megoldását igénylik, vagy a duál feladatot használják fel.

Először megmutatjuk, hogy a lineáris programozás elméletének felhasználásával hogyan állíthatjuk elő a hátizsák feladat optimális megoldását. Ezután újra belátjuk most már elemi eszközökkel, hogy a kapott megoldás optimális.

Tekintsünk egy általános felső korlátos lineáris programozási feladatot:

$$\begin{aligned} & \max c'x \\ (2.1) \quad & Ax \leq b \\ & 0 \leq x \leq U. \end{aligned}$$

Jelöljük a_j -vel az A mátrix j -edik oszlopát.

A következő tétel egyszerű következménye a kiegészítő eltérések gyenge tételének [28].

2.1. TÉTEL (felső korlátos LP optimális megoldása).

Tekintsük a (2.1) feladat egy x megengedett megoldását. Az x vektor akkor és csak akkor optimális megoldása a (2.1) feladatnak, ha létezik egy olyan $q > 0$ vektor, amelyre

1. $(b - Ax)'q = 0$,
2. $c_j - q'a_j > 0$ esetén $x_j = U_j$,
3. $c_j - q'a_j < 0$ esetén $x_j = 0$.

A 2.1. tétel lehetővé teszi az (1.2) folytonos hátizsák feladat optimális megoldásának explicit felírását. A megoldást jelöljük x^c -vel, az optimumot z_c -vel. Tegyük fel továbbá, hogy

$$(2.2) \quad c \geq 0 \quad \text{és} \quad a > 0.$$

Ha valamely indexre

$$c_j < 0 \quad \text{és} \quad a_j < 0,$$

akkor az

$$x'_j = 1 - x_j$$

helyettesítéssel feltevésünk kielégíthető.

Ha

$$c_j < 0 \quad \text{és} \quad a_j > 0,$$

akkor

$$x_j^c = 0,$$

ha viszont

$$c_j > 0 \quad \text{és} \quad a_j < 0,$$

akkor

$$x_j^c = 1.$$

(Az előbbi állítás nyilvánvaló, de a 2.1. tételből is következik.)

A fenti feltevés így nem jelenti az általánosság megszorítását, csupán a tárgyalás egyszerűsítése miatt célszerű.

Vezessük be a következő jelöléseket:

$$\begin{aligned} q_j &= c_j/a_j; \quad j=1, \dots, n, \\ N &= \{1, \dots, n\}, \\ E_k &= \{j \in N : q_j = q_k\}; \quad k=1, \dots, n, \\ U_k &= \{j \in N : q_j > q_k\}; \quad k=1, \dots, n, \\ G_k &= U_k + \{j \in E_k : j \leq k\}; \quad k=1, \dots, n, \\ \bar{G}_k &= N - G_k. \end{aligned}$$

2.1. KÖVETKEZMÉNY (folytonos hátizsák feladat optimális megoldása).

Tekintsük az (1.2) feladatot a (2.2) feltevés figyelembevételével. Rendezzük a változókat a q_j hányadosok csökkenő sorrendje szerint! Ekkor a feladat opti-

mális megoldását x^c -vel jelölve:

$$(2.3) \quad x_j^c = \begin{cases} 1, & j \in G_p - \{p\}, \\ (1/a_p)(b - \sum_{j \in G_p - \{p\}} a_j), & \text{ahol } p = \min \{k: \sum_{j \in G_k} a_j > b\}, \\ 0, & j \in \bar{G}_p. \end{cases}$$

Bizonyítás. A 2.1. tétel szerint ahhoz, hogy az x megengedett megoldás optimális legyen, léteznie kell (és ez elegendő) olyan q számnak, amelyre:

$$c_j/a_j > q \text{ esetén } x_j = 1,$$

$$c_j/a_j < q \text{ esetén } x_j = 0,$$

$$\text{és } (b - a'x)q = 0.$$

Megjegyzés. A (2.3) megoldás optimalitását közvetlenül is beláthatjuk. Legyen x egy tetszőleges megengedett megoldás. Ekkor

$$\begin{aligned} & \sum_{j \in N} c_j x_j^c - \sum_{j \in N} c_j x_j = \sum_{j \in N} c_j (x_j^c - x_j) = \\ &= \sum_{\{j: a_j > a_p\}} c_j (x_j^c - x_j) + \sum_{\{j: a_j = a_p\}} c_j (x_j^c - x_j) + \sum_{\{j: a_j < a_p\}} c_j (x_j^c - x_j) \cong \\ &\cong \sum_{\{j: c_j > \frac{c_p}{a_p} a_j\}} a_j (c_p/a_p) (x_j^c - x_j) + \sum_{\{j: c_j = \frac{c_p}{a_p} a_j\}} a_j (c_p/a_p) (x_j^c - x_j) + \\ &+ \sum_{\{j: c_j < \frac{c_p}{a_p} a_j\}} a_j (c_p/a_p) (x_j^c - x_j) = (c_p/a_p) \left(\sum_{j \in N} a_j x_j^c - \sum_{j \in N} a_j x_j \right) \cong 0. \end{aligned}$$

Az első alsó becslés $(x_j^c - x_j)$ előjele miatt igaz.

3. A bináris feladat

Viszonylagos egyszerűsége mellett a hátizsák feladat megoldásánál már felmerül a diszkrét programozásban előforduló nehézségek egy jelentős része. Ennél fogva a megoldási módszerek alapja megegyezik az általánosabb diszkrét problémákéval. Miért szükséges mégis e speciális feladat behatóbb vizsgálata? Egyrészt azért, mert az általánosaknál lényegesen hatékonyabb algoritmusokat lehet kidolgozni a megoldására, másrészt az előbbieik alapján eredményesen használható bonyolultabb problémák megoldása során részfeladatként, vagy relaxációként. Ezekre a kérdésekre a későbbiekben még visszatérünk.

A diszkrét programozási problémák megoldására ismert legelső módszer, amelyet a hátizsák feladatra alkalmaztak, a dinamikus programozási elv felhasználásán alapult ([6]). Az ilyen algoritmusok hatékonysága azonban a feladat méretén (a változók számán) kívül a b jobb oldal nagyságától is lineárisan függ. Ezért használatuk csak bizonyos esetekben célszerű, így egyes darabolási problémák megoldása során, amikor nagyszámú hátizsák feladatot kell megoldanunk különböző jobb oldalakkal. GILMORE és GOMORY ilyen irányú munkáikban [12, 13, 14, 15] mutat-

ták meg, hogy a hátizsák feladat speciális tulajdonságait hogyan lehet az általánosanál jobb dinamikus programozási algoritmus kidolgozásához felhasználni.

Általában nagyméretű 0—1-es hátizsák feladatok megoldására a fabejáró algoritmusok bizonyultak legalkalmasabbnak. Az első ilyen eljárás LAND és DOIG [22] munkájában található. A korlátozás és szétválasztásnak nevezett módszert KOLESAR [20] alkalmazta a hátizsák feladatra. Hatékonysága a folytonos optimum gyors kiszámíthatóságának köszönhető. A következő jelentős lépés GREENBERG és HEGERICH [16] algoritmus, mely szintén a folytonos feladatot használja fel. Ezután merült fel a méretcsökkentés gondolata ([18]). A *Lagrange relaxáció*s, vagy büntetési méretcsökkentés általánosabb megfogalmazásban már GEOFFRION korábbi munkáiban is megtalálható [11], míg a hátizsák feladatra FAYARD és PLATEAU alkalmazta ([8]).

A legutóbbi időkben még számos algoritmus született, amelyekkel egyre hatékonyabban és egyre nagyobb méretű feladatokat lehet megoldani. Ezek lényegében a fent említett eljárások ügyes módosításai és kombinációi.

Meg kell említenünk, hogy a méretcsökkentési módszerek egy esetben használhatatlanok. Ez az azonos hányadosú hátizsák probléma, ahol a c_j/a_j hányadosok minden j -re megegyeznek. Ilyenkor egyébként a fabejáró algoritmusok is csődöt mondanak, azaz csaknem teljes leszámllálást valósítanak meg ([1, 19]). Arra az esetre, amikor a jobb oldal viszonylag nagy értéke miatt a dinamikus programozás sem alkalmazható, az [1] dolgozatban találhatunk megközelítést. A továbbiakban nem ilyen feladatokkal foglalkozunk.

a) Fabejáró algoritmusok

KOLESAR 1967-ben közölt algoritmus [20] immár klasszikussá vált, sőt az oktatásban magának a korlátozás és szétválasztás módszerének illusztrálására szokták használni ([21]). Memóriaigénye és viszonylagos lassúsága miatt azonban már ritkán alkalmazzák.

Az azóta megjelent algoritmusok csaknem mindegyike a GREENBERG és HEGERICH által 1970-ben javasolt leszámllálási sémát alkalmazta a feladat egzakt megoldására.

A leszámllálási módszerek általános váza KOVÁCS LÁSZLÓ BÉLA [21] könyvében található meg.

Ennek lényege, hogy a változókat beválasztva, illetve kizárva felépíti azok lehetséges részhalmazait, és megvizsgálja, hogy egy részhalmaz továbbépítésével kielégíthetők-e a feltételek, illetve javítható-e a célfüggvény.

Egy változó beválasztását 1, kizárását 0 értékű rögzítésnek nevezzük. A lekötés és következmény fogalmát szintén a [21] könyvnek megfelelően használjuk.

Az időrendi sorrendtől eltérően először HOROWITZ és SAHNI algoritmusát [17] írjuk le. Ez amellett, hogy szintén a GREENBERG és HEGERICH féle gondolatokon alapszik, egy szokásos értelemben vett leszámllálási eljárás.

Vezessük be a következő jelöléseket az eddigieket kiegészítve:

$R_1 := \{j \in N : \text{az } x_j \text{ változót 1 értéken rögzítettük}\},$

$R_0 := \{j \in N : \text{az } x_j \text{ változót 0 értéken rögzítettük}\},$

$F := \{j \in N : \text{az } x_j \text{ változót nem rögzítettük}\},$

$L_{R_1} := \text{az } R_1\text{-hez tartozó indexek közül az utoljára beválasztott},$

$L_{R_0} := \text{az } R_0\text{-hoz tartozó indexek közül az utoljára beválasztott}.$

$H \subset N$ esetén

$$b_H := b - \sum_{j \in H} a_j,$$

$$z_H := \sum_{j \in H} c_j.$$

z_L : egy megengedett (az eljárás során talált legjobb) célfüggvényérték.

Nyilvánvaló, hogy

$$N = R_1 + R_0 + F.$$

$T \subset N$ egy ideiglenes munkahalmat jelöl. Az algoritmus a következő:

1. Rendezzük a változókat a $q_j = c_j/a_j$ számok csökkenő sorrendje szerint, és legyen $j := 1$, $z_L := 0$, $T := R_1 := R_0 := \emptyset$, $I := N + \{n+1\}$, $F := I$.

2. Mindaddig, amíg

$$b_{R_1+T} - a_j \geq 0 \quad \text{és} \quad j \leq n,$$

legyen

$$T := T + \{j\},$$

és

$$j := \min \{i : i > j \quad \text{és} \quad i \in F\}.$$

3. Ha $j \geq n$, akkor menjünk az 5. lépésre.

4. Következménykeresés. Ha $j < n$, akkor vizsgáljuk meg, hogy teljesül-e a következő egyenlőtlenség:

$$z_{R_1+T} + q_j b_{R_1+T} > z_L.$$

Ha igen, akkor

$$R_1 := R_1 + T, \quad T := \emptyset,$$

$$R_0 := R_0 + \{j\},$$

$$F := I - (R_0 + R_1),$$

$$j := \min \{i : i > j \quad \text{és} \quad i \in F\},$$

Ha nem, akkor menjünk a 6. lépésre.

5. Az optimum javítása. Ha

$$z_L < z_{R_1+T},$$

akkor

$$z_L := z_{R_1+T}.$$

6. Ágcsere.

$$j := L_{R_1}$$

$$R_1 := R_1 - \{j\},$$

$$R_0 := R_0 + \{j\} - \{i : i \in R_0 \quad \text{és} \quad i > j\},$$

$$F := I - (R_0 + R_1),$$

$$j := \min \{i : i > j \quad \text{és} \quad i \in F\},$$

menjünk a 2. lépésre.

A fenti algoritmust a különböző rögzítések melletti folytonos optimumnak a leszámplálási eljárás menetébe illő könnyű kiszámíthatósága és az ennek segítségével gyorsan ellenőrizhető teszt teszi hatékonnyá.

GREENBERG és HEGERICH algoritmusának [16] lényegében ugyanezek az előnyei, csak a leszámplálást végzi más, jól megindokolható módon. A Horowitz—Sahni-eljárásban a 0 rögzítések vagy következmények, vagy lekötések. Habár a 0 rögzítések itt is 1 rögzítések következményeiként állnak elő, a leszámplálás az 1 rögzítéseket nem veszi figyelembe, viszont a következményként kapott 0 lekötéseket sima rögzítésként kezeli, és csak ezeken hajt végre ágcserét.

Greenberg és Hegerich algoritmus a következő módon írható le:

1. Rendezzük a változókat a $q_j = c_j/a_j$ számok csökkenő sorrendje szerint, és legyen $j := 1$, $z_L := 0$, $T := R_0 := R_1 := \emptyset$, $I := N + \{n + 1\}$, $F := I$.

2. Mindaddig, amíg

$$b_{R_1+T} - a_j \geq 0 \quad \text{és} \quad j \leq n,$$

legyen

$$T := T + \{j\},$$

és

$$j := \min \{i : i > j \quad \text{és} \quad i \in F\}.$$

3. Ha $j \leq n$, akkor menjünk az 5. lépésre.

4. Következménykeresés. Ha $j < n$, akkor vizsgáljuk meg, hogy teljesül-e a következő egyenlőtlenség:

$$z_{R_1+T} + q_j b_{R_1+T} > z_L.$$

Ha igen, akkor

$$R_0 := R_0 + \{j\},$$

$$F := I - (R_0 + R_1),$$

$$j := \min \{i : i > j \quad \text{és} \quad i \in F\},$$

menjünk a 2. lépésre.

Ha nem, akkor menjünk a 6. lépésre.

5. Az optimum javítása. Ha

akkor

$$z_L < z_{R_1+T},$$

$$z_L := z_{R_1+T}.$$

6. Ágcseré.

$$j := L_{R_0},$$

$$R_0 := R_0 - \{j\},$$

$$R_1 := R_1 + \{j\},$$

$$T := \emptyset,$$

$$j := \min \{F\},$$

menjünk a 2. lépésre.

b) *A méretcsökkentés*

Bár az eddig leírt eljárásokkal elért számítógépes eredmények is jók voltak, mégis sikerült még jobb módszereket kidolgozni. Ezek kivétel nélkül felhasználják a méretcsökkentés gondolatát. Ennek lényege, hogy azokat a változókat, amelyeket lehetséges, még a költségesebb leszámítás előtt lerögzítjük az optimális értékükre. Ez azért előnyös, mert a leszámítás sebessége legrosszabb esetben exponenciális függvénye a benne résztvevő változók számának, míg a méretcsökkentés polinomiális, még ha esetleg csekély hatású is.

Egészítsük ki az eddigi jelöléseket:

$z_u(x_j = \varepsilon)$: a diszkrét feladat optimális célfüggvényértékének egy felső becslése az x_j változó ε rögzítése mellett.

X_0 , illetve X_1 : a 0-ra, illetve 1-re rögzített változók indexeinek halmaza.

A méretcsökkentési algoritmusok egy változót akkor rögzítenek le, ha az ellenkező értékű rögzítés mellett a célfüggvény egy felső becslése rosszabb, mint egy tetszőleges alsó becslés.

A méretcsökkentési algoritmusok a következő keretbe foglalhatók:

1. $X_0 := X_1 := \emptyset$,
2. $\varepsilon_k := \begin{cases} 0, & \text{ha } k \in G_p, \\ 1, & \text{ha } k \in \bar{G}_p + E_p. \end{cases}$
3. $k := 1, \dots, n$ értékekre,
 ha $z_u(x_k = \varepsilon_k) < z_L$,
 akkor $x_k := 1 - \varepsilon_k$,
 $X_{1-\varepsilon_k} := X_{1-\varepsilon_k} \cup \{k\}$.

Az ismert módszerek két osztályba sorolhatók:

a) INGARGIOLA és KORSH módszerének továbbfejlesztései (nevezzük ezeket 1-es módszernek).

b) *Lagrange-féle relaxációs méretcsökkentés* (nevezzük ezeket 2-es módszernek). Ezek lényegében az alsó becslés kiszámításának módjában és a felső becslés megválasztásában különböznek.

Az 1-es módszer [18] időben az első méretcsökkentési eljárás. Itt $z_u(x_k = \varepsilon)$ az $x_k = \varepsilon$ rögzítés melletti folytonos optimum, amely mint láttuk könnyen kiszámítható. Ennek az algoritmusnak jelentős előnye még, hogy az alsó becslést is növelni tudja. Ugyanis ha az általános keretben szereplő feltétel nem teljesül, vagyis

$$z_u(x_k = \varepsilon_k) > z_L,$$

akkor és csak akkor van esély arra, hogy a $z_u(x_k = \varepsilon_k)$ kiszámításához vezető úgynevezett mohó algoritmust folytatva z_L becslésnél jobb megengedett célfüggvényértékeket kapunk.

A legegyszerűbb mohó algoritmus a következőképpen írható le:

1. $z_L(x_k = \varepsilon_k) := \varepsilon_k c_k$, $b' := b - \varepsilon_k a_k$, $j := 0$.
2. Ha $j < n$, akkor $j := j + 1$, egyébként STOP.

3. Ha $a_j \leq b'$ és $j \neq k$,
 akkor $z_{L_1}(x_k = \varepsilon_k) := z_{L_1}(x_k = \varepsilon_k) + c_j$,
 $b' := b' - a_j$,
 ha $b' < \min \{a_j\}$, akkor STOP.

Menjünk a 2. lépésre.

Ezután az általános keretet a következőkkel lehet kiegészíteni:

- Ha $z_u(x_k = \varepsilon_k) > z_L$,
 akkor számítsuk ki a $z_{L_1}(x_k = \varepsilon_k)$ értéket.
 — Ha $z_{L_1}(x_k = \varepsilon_k) > z_L$,
 akkor $z_L := z_{L_1}(x_k = \varepsilon_k)$.

Most vizsgáljuk meg a *Lagrange-féle relaxációs eljárást* (2-es módszer).

A következőkben a tárgyalás ott, ahol lehet, általános lesz (több feltételes feladat), de a hátizsák feladatra alkalmazás mindenütt triviális.

Tekintsük a következő lineáris diszkrét feladatot, amelynek folytonos megfelelője a (2.1) feladat $U=1$ esetén.

$$(3.1) \quad \begin{aligned} &\max \mathbf{c}'\mathbf{x} \\ &\mathbf{Ax} \leq \mathbf{b} \\ &\mathbf{x} \in \{0, 1\}^n \end{aligned}$$

Jelöljük itt is \mathbf{a}_j -vel az \mathbf{A} mátrix j -edik oszlopvektorát.

Legyen z_L (3.1) egy megengedett célfüggvényértéke, \mathbf{x}^D pedig az optimális megoldása.

Legyen $\mathbf{q} > \mathbf{0}$ és

$$M(\mathbf{q}) = \max \{\mathbf{c}'\mathbf{x} + \mathbf{q}'(\mathbf{b} - \mathbf{Ax}) : \mathbf{x} \in \{0, 1\}^n\}$$

Könnyen látható, hogy

$$(3.2) \quad M(\mathbf{q}) = \max \{\mathbf{c}'\mathbf{x} + \mathbf{q}'(\mathbf{b} - \mathbf{Ax}) : \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\} = \mathbf{q}'\mathbf{b} + (\mathbf{c}' - \mathbf{q}'\mathbf{A})^+ \mathbf{1}.$$

($d^+ = \max \{0, d\}$ a szokásos értelmezés.)

3.1. TÉTEL. Legyen $\mathbf{q} \geq \mathbf{0}$. Ha

$$\begin{aligned} &\text{akkor} \\ &\quad M(\mathbf{q}) - |c_j - \mathbf{q}'\mathbf{a}_j| < z_L, \\ &\quad c_j - \mathbf{q}'\mathbf{a}_j < 0 \quad \text{esetén} \quad x_j^D = 0, \\ &\quad c_j - \mathbf{q}'\mathbf{a}_j > 0 \quad \text{esetén} \quad x_j^D = 1. \end{aligned}$$

Bizonyítás. (3.2) alapján látható, hogy

$$\begin{aligned} &\text{ahol} \\ &\quad M(\mathbf{q}) - |c_j - \mathbf{q}'\mathbf{a}_j| = \max \{\mathbf{c}'\mathbf{x} + \mathbf{q}'(\mathbf{b} - \mathbf{Ax}) : \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, x_j = \varepsilon_j\}, \\ &\quad \varepsilon_j = 0, \quad \text{ha} \quad c_j - \mathbf{q}'\mathbf{a}_j > 0, \\ &\quad \varepsilon_j = 1, \quad \text{ha} \quad c_j - \mathbf{q}'\mathbf{a}_j < 0, \end{aligned}$$

továbbá

$$\begin{aligned}
 & \max \{ \mathbf{c}'\mathbf{x} + \mathbf{q}'(\mathbf{b} - \mathbf{A}\mathbf{x}) : \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \quad x_j = \varepsilon_j \} \cong \\
 (3.3) \quad & \cong \max \{ \mathbf{c}'\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \quad x_j = \varepsilon_j \} \cong \\
 & \cong \max \{ \mathbf{c}'\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \in \{0, 1\}^n, \quad x_j = \varepsilon_j \}.
 \end{aligned}$$

Tehát $M(\mathbf{q}) - |c_j - \mathbf{q}'\mathbf{a}_j|$ felső becslése a (3.1) feladat $x_j = \varepsilon_j$ rögzítés melletti optimális célfüggvény értékének. Innen azonnal adódik a tétel állítása.

Most az általános tárgyalást megszakítjuk, néhány idevonatkozó algoritmikus észrevétel kedvéért.

3.1. Megjegyzés. Az 1-es módszer legalább annyi változó optimális értékét határozza meg, mint a 2-es, vagyis INGARGIOLA és KORSH módszere erősebb.

Bizonyítás. (3.3) alapján látható, hogy $M(\mathbf{q}) - |c_j - \mathbf{q}'\mathbf{a}_j|$ nemcsak az $x_j = \varepsilon_j$ rögzítés melletti diszkrét optimumnak, de a (3.1) feladat ($x_j = \varepsilon_j$ rögzítés melletti) folytonos optimumának is felső becslése. Márpedig az 1-es módszer a hátizsák feladat esetén pontosan ez utóbbit használja.

A 3.1. megjegyzés felveti azt a kérdést, hogy miért van szükség a 2-es módszerre. A válasz az, hogy ez az eljárás lényegesen gyorsabb, mint az 1-es. Az 1-es módszer műveletszáma a változók számának négyzetével, míg a 2-es módszeré a változók számával arányos.

Most foglaljuk össze az 1-es, illetve a 2-es módszer előnyeit. Az 1-es módszer erősebb és állandóan növeli az alsó becslést, amire a 2-es módszerben is szükség van. A 2-es módszer előnye a gyorsaság. Ezek az előnyök azonban egyesíthetők [3], ugyanis az 1-es módszer alkalmazása során mindig rendelkezésünkre állnak a 2-es módszer tesztjéhez szükséges adatok. Így mielőtt kiszámítanánk egy adott rögzítés melletti folytonos optimumot (ami lassítja az 1-es módszert) megvizsgáljuk, hogy a 2-es módszer tesztje eredményes-e. Ha igen, akkor nincs szükség további számításokra. (Ekkor nyilván a z_L alsó becslés sem növelhető.)

Most visszatérünk az általános tárgyaláshoz, hogy megvizsgáljuk a \mathbf{q} Lagrange-szorzó vektor megválasztásának kérdését.

A 3.1. tételből kiindulva a legjobb Lagrange-szorzó vektort úgy kellene definiálni, hogy segítségével a lehető legtöbb változó legyen rögzíthető. Tehát egy egyenlőtlenség rendszerhez egy olyan nem negatív „megoldást” kellene találni, amely a lehető legtöbb egyenlőtlenséget elégíti ki. E feladat megoldása azonban olyan költséges lenne, hogy értelmetlenné tenné magát a mértécsökkentést.

Ezek után nevezzük legjobbnak azt a \mathbf{q}^* Lagrange-szorzó vektort, amelyre

$$(3.4) \quad M(\mathbf{q}^*) = \min_{\mathbf{q} \geq \mathbf{0}} M(\mathbf{q}).$$

Vegyük észre, hogy (3.4) a (3.1) feladat folytonos megfelelőjének duáljával ekvivalens. Ez a duál ugyanis a következő:

$$\begin{aligned}
 & \min \mathbf{q}'\mathbf{b} + \mathbf{t}'\mathbf{1} \\
 (3.5) \quad & \mathbf{q}'\mathbf{A} + \mathbf{t}' \cong \mathbf{c}' \\
 & \mathbf{q} \geq \mathbf{0} \quad \text{és} \quad \mathbf{t} \geq \mathbf{0}.
 \end{aligned}$$

Ebből a t változó vektort kiküszöbölve a következő ekvivalens feladatot kapjuk:

$$\min_{q \geq 0} \{q'b + (c' - q'A) + 1\}.$$

Ez viszont (3.2) szerint pontosan a (3.4) feladat.

Ezzel beláttuk a következő tételt.

3.2. TÉTEL. A fenti (3.4) értelemben vett legjobb *Lagrange-szorzó vektor* pontosan a folytonos feladat duáljának optimális megoldása.

Megjegyzés: A folytonos feladat duáljának optimális megoldását a diszkrét programozásban s -feltételek generálásához is szokták használni [21]. A bizonyos értelemben legjobb s -feltételhez szükséges szorzókat a következő feladat megoldása adja:

$$(3.6) \quad \min_{q \geq 0} \max_{0 \leq x \leq 1} \{c'x : q'Ax \leq q'b\}.$$

A belső feladat duálját felírva és a két minimumot összevonva pontosan a (3.5) feladatot kapjuk. Tehát valójában a (3.4) és a (3.6) feladat ugyanaz, azaz

$$\min_{q \geq 0} \max_{0 \leq x \leq 1} \{c'x + q'(b - Ax)\} = \min_{q \geq 0} \max_{0 \leq x \leq 1} \{c'x : q'Ax \leq q'b\}.$$

Visszatérve ismét a hátizsák feladathoz a 3.2. tétel, a 2.1. tétel és a 2.1. következmény felhasználásával a legjobb *Lagrange-szorzó*

$$q^* = q_p = c_p/a_p,$$

ahol p a 2.1. következményben van definiálva.

A q_p szám optimalitása elemi eszközökkel is belátható. A 2.1. következmény alapján a folytonos optimum (z_c) pontosan

$$z_c = M(q_p) = \sum_{j=1}^{p-1} c_j + \frac{c_p}{a_p} \left(b - \sum_{j=1}^{p-1} a_j \right).$$

Ezután belátjuk, hogy minden $q \geq 0$ -ra

$$z_c = M(q_p) \leq M(q).$$

Két esetet különböztetünk meg, és mindkettőben adunk egy z_c -nél jobb megoldást.

Ha $q_p \leq p$, akkor legyen

$$x_j = \begin{cases} 1, & j = 1, \dots, p-1, \\ 0, & j = p, \dots, n. \end{cases}$$

$$c'x + q(b - a'x) = \sum_{j=1}^{p-1} c_j + q \left(b - \sum_{j=1}^{p-1} a_j \right) \geq z_c, \text{ mivel } b - \sum_{j=1}^{p-1} a_j \geq 0.$$

Ha $q \leq q_p$, akkor legyen

$$x_j = \begin{cases} 1, & j = 1, \dots, p. \\ 0, & j = p+1, \dots, n. \end{cases}$$

$$\begin{aligned} c'x + q(b - a'x) &= \sum_{j=1}^p c_j + q \left(b - \sum_{j=1}^p a_j \right) = \\ &= \sum_{j=1}^{p-1} c_j + q \left(b - \sum_{j=1}^{p-1} a_j \right) + c_p - qa_p \cong \\ &\cong \sum_{j=1}^{p-1} c_j + qa_p \frac{1}{a_p} \left(b - \sum_{j=1}^{p-1} a_j \right) + (c_p - qa_p) \frac{1}{a_p} \left(b - \sum_{j=1}^{p-1} a_j \right) = \\ &= z_c = \sum_{j=1}^{p-1} c_j + c_p \frac{1}{a_p} \left(b - \sum_{j=1}^{p-1} a_j \right). \end{aligned}$$

mivel

$$c_p - qa_p \cong 0 \quad \text{és} \quad \frac{1}{a_p} \left(b - \sum_{j=1}^{p-1} a_j \right) \leq 1.$$

A fentiek alapján a *Lagrange-féle relaxációs méretcsökkentést* célszerű a q_p szorzó segítségével végezni. Ezt azonban nemcsak az elméleti megfontolások, hanem a számítógépes tapasztalatok is igazolják. Sőt a méretcsökkentés hatékonysága általában rendkívül érzékeny a szorzó jó megválasztására.

4. A rendezés problémája

Az előző fejezetekből többek között az tűnik ki, hogy a hátizsák feladat változóinak a q_j hányadosok csökkenő sorrendje szerinti rendezése a megoldó és méretcsökkentő algoritmusok hatékonyságának érdekében lényeges.

Ez egy ideig nem okozott problémát. A 70-es évek második felében azonban a méretcsökkentési eljárások olyan hatékonyá váltak, hogy a változók kezdeti rendezési idejének és a teljes megoldás idejének aránya 50% fölé nőtt [2, 4, 9].

A megoldási módszereknek, valamint INGARGIOLA és KORSH méretcsökkentési módszerének lényegéhez tartozik a változók rendezettsége.

A *Lagrange-féle relaxációs méretcsökkentés* esetén a rendezés szükséges egyrészt a legjobb szorzó megtalálásához, másrészt célszerű egy jó alsó becslés eléréséhez használt heurisztikában.

Ebben a fejezetben olyan módszerekkel foglalkozunk, amelyek felhasználásával a *Lagrange-féle relaxációs méretcsökkentés* esetén kiküszöbölhető a változók teljes rendezése. Az így kapott eljárás nagyméretű feladatok esetén jelentősen csökkenti a megoldás idejét, ha beszámítjuk a rendezési időket is.

a) *A becsült optimum módszer* [4]

Ebben a szakaszban a *Lagrange-féle relaxációs méretcsökkentés*hez használt becslés a szokásostól eltérően nem egy hosszadalmas és rendezés után keresett megengedett megoldáshoz tartozó célfüggvényérték lesz. Ehelyett a diszkrét optimum egy \bar{z} becslését használjuk.

Ha ez a becslés kisebb a diszkrét optimumnál, akkor a segítségével végrehajtott méretcsökkentés nyilván szabályos. Ekkor a maradék feladatot megoldva megkapjuk az optimális megoldást. Ha a becslés nagyobb a diszkrét optimumnál, akkor a méretcsökkentés különösen nagy mértékű. Ezzel szemben a fennmaradó viszonylag kis feladat megoldásával nem biztos, hogy megkapjuk a diszkrét optimumot (ui. szabálytalan rögzítések is lehetnek), sőt e feladatnak lehet, hogy megengedett megoldása sincs.

Mégis megadható egy olyan eljárás, amely a fent említett becsléseknek egy sorozatán alapszik és végül biztosan megadja az optimumot. A módszer előnye abban rejlik, hogy a becslések végig az optimum közelében maradnak, így az azokkal végrehajtott méretcsökkentések után megoldandó feladatok kicsik lesznek. Most térjünk rá a konkrét algoritmusra, amelyet „becsült optimum módszernek” nevezhetünk.

Becsült optimum algoritmus

1. $i := 1$, becsüljük meg a (3.1) feladat diszkrét optimális célfüggvényértékét! (legyen ez \bar{z}_i).

2. Csökkentsük a feladat méretét *Lagrange-féle relaxációval* a z_L alsó becslés helyett \bar{z}_i felhasználásával és egy tetszőleges *Lagrange-szorzóval* (3.1. tétel).

3. Ha a fennmaradó feladatnak nincs megengedett megoldása, akkor menjünk a 6. lépésre.

4. Egyébként oldjuk meg a fennmaradó kisméretű feladatot! (Legyen z_{R_i} az optimális célfüggvényérték!)

5. Ha

$$z_{R_i} \cong \bar{z}_i.$$

akkor megtaláltuk az optimális megoldást és z_{R_i} az optimum. STOP!

6. Egyébként módosítsuk a \bar{z}_i becslést egy megfelelő ε értékkel!

$$\bar{z}_{i+1} := \bar{z}_i - \varepsilon_i,$$

ahol

$$\varepsilon_i > 0,$$

és

$$\sum_{i=1}^{\infty} \varepsilon_i > \bar{z}_1 \quad i := i + 1.$$

Menjünk a 2. lépésre.

A fenti algoritmus végességének igazolásához elég azt belátnunk, hogy véges sok iteráció után teljesül a

$$z_{R_i} \cong \bar{z}_i$$

egyenlőtlenség. Ezt azonban az ε_i sorozatra tett feltevések és a következő egyenlőtlenségek biztosítják:

$$z_{R_1} \leq z_{R_2} \leq \dots \leq z_{R_i}, \\ \bar{z}_1 > \bar{z}_2 > \dots > \bar{z}_i.$$

Az algoritmus hatékonysága erősen függ a \bar{z}_1 kezdeti becsléstől és ε_i iterációnkénti megválasztásától.

A becslült optimum módszer alapgondolata általánosabban is megfogalmazható.

Tekintsünk egy olyan problémát, amelyhez ismerünk egy megoldási módszert. Tételezzük fel, hogy bizonyos mennyiségi információ birtokában lényegesen lecsökkenthetnénk a megoldási időt, ez az információ azonban csak éppen e megoldás után kerül birtokunkba.

Ebben az esetben célszerű lehet becsléshez folyamodni, és segítségével megoldani a feladatot. Ha az ezután végül megkapott információ a becslésünktől függetlenül alátámasztja annak helyességét, akkor a feladatot meg is oldottuk. Ellenkező esetben az eljárást ismételni kell módosított becslésekkel.

A hátizsák feladat esetén a módosításra szabott egyszerű feltételekkel biztosítani tudtuk a módszer végességét.

b) A folytonos optimum meghatározása teljes rendezés nélkül

Az előző pontban a *Lagrange-féle relaxációs méretcsökkentés* egyik kritikus adatával foglalkoztunk. Ebben a pontban egy másik ilyen adat, a legjobb *Lagrange-szorzó* meghatározása lesz a cél.

A 2.1. következmény szerint célunk elérésének legegyszerűbb eszköze a változók rendezése és a „hátizsák betöltése”. Most azonban olyan módszereket keresünk, amelyekben erre nincs szükség.

A következő algoritmus egyszerűsített, de a lényeget tartalmazó formája annak, amit a FAYARD és PLATEAU, valamint a BALAS és ZEMEL szerzőpárosok feltehetően egymástól függetlenül írtak le 1977-ben [2, 9]. Ennek alapgondolata egy „quicksort” típusú rendezés, amelyet azonban a pivot megtalálása után szükségtelen befejezni. Az algoritmus átlagos lépésszáma a változók számának lineáris függvénye.

QUICKSORT algoritmus. (A jelölések megfelelnek a 2. szakaszban bevezetetteknek.)

0. $K := \{1, \dots, n\}$, $U := \emptyset$.

1. Válasszuk ki K egy elemét ($k \in K$)!

2. Ha

$$\sum_{j \in U_k \wedge K} a_j > b,$$

akkor $K := U_k \wedge K$, és menjünk az 1. lépésre.

3. Egyébként, ha

$$\sum_{j \in (U_k + E_k) \wedge K} a_j > b$$

akkor $q^* := q_k$, STOP!

4. Egyébként

$$b := b - \sum_{j \in (U_k + E_k) \wedge K} a_j,$$

$$K := K - (U_k + E_k),$$

menjünk az 1. lépésre.

A továbbiakban egy olyan módszerről lesz szó [4], amelynek célja ugyanaz, mint az előbbié, viszont az előző szakaszban leírt becsült optimum módszerhez illeszkedik. Az algoritmus leírása előtt be kell látnunk néhány lemmát.

Jelölje \bar{q} a q^* szorzó egy becslését. A 3.1. tételben a \bar{q} szorzót és egy tetszőleges becslést felhasználva a következő halmazokat állíthatjuk elő:

$$\begin{aligned} R_1 &:= \{j \in N: \text{az } x_j \text{ változót } 1 \text{ értéken rögzítettük}\}, \\ R_0 &:= \{j \in N: \text{az } x_j \text{ változót } 0 \text{ értéken rögzítettük}\}, \\ F &:= \{j \in N: \text{az } x_j \text{ változót nem rögzítettük}\}. \end{aligned}$$

Nyilvánvaló, hogy

$$N = R_0 + R_1 + F.$$

Jelöljük r -rel a fennmaradó feladat pivot indexét. (Feltételezzük, hogy a fennmaradó feladatnak van megoldása, azaz $\sum_{j \in R_1} a_j \leq b$.)

A továbbiakban a változóknak a q_j számok csökkenő sorrendje szerinti rendezetlenségét csak a bizonyításokhoz szükséges definíciókban tételezzük fel.

Tehát:

$$r := \min \{k \in F: \sum_{j \in R_1 + (F \wedge G_k)} a_j > b\}.$$

Ha a fent szereplő halmaz üres, akkor a következő lemmák $r := n+1$ és $q_{n+1} := 0$ felvételével nyilvánvalók, így ezzel az esettel külön nem foglalkozunk.

4.1. LEMMA.

$$r = \min \{k \in N: \sum_{j \in R_1 + (F \wedge G_k)} a_j > b\}.$$

Bizonyítás. Legyen

$$r' := \min \{k \in N: \sum_{j \in R_1 + (F \wedge G_k)} a_j > b\}!$$

Belátjuk, hogy $r' \in F$, amiből már következik, hogy $r = r'$. Ha $r' \notin F$, akkor $r' \in R_0 + R_1$. De ekkor legyen

$$r'' := \max \{j \in FG_{r'}\}!$$

Ha $F \wedge G_{r'} = \emptyset$, akkor r' definíciójából

$$\sum_{j \in R_1} a_j > b,$$

ami ellentmond a fennmaradó feladat megengedettségének, amit viszont feltételeztünk. Ha $F \wedge G_{r'} \neq \emptyset$, akkor $r'' < r'$ és

$$R_1 + (F \wedge G_{r''}) = R_1 + (F \wedge G_{r'}),$$

ami ellentmond r' definíciójának.

Ezekután legyen

$$h := \max \{k \in N: q_k > \bar{q}\}!$$

Ha a fent szereplő halmaz üres, akkor legyen $h := 0$ és $G_0 := \emptyset$!

A következő lemma tartalmazza a q^* szorzó megtalálására szolgáló új módszer alap gondolatát.

4.2. LEMMA. A legjobb $q^* = q_p$ szorzó egy tetszőleges \bar{q} szorzó és az ennek segítségével elvégzett méretcsökkentés után fennmaradó feladat legjobb q_r szorzója közé esik, azaz ha

$$q' = \max \{\bar{q}, q_r\} \quad \text{és} \quad q'' = \min \{\bar{q}, q_r\}$$

akkor

$$q' \cong q^* = q_p \cong q''$$

Bizonyítás. A bizonyítást három esetre választjuk szét.

1. eset: $\bar{q} > q_p$.

Ekkor $G_h \subset G_p - \{p\}$.

Legyen k az az index, amelyre $G_k = G_p - \{p\}$. Mivel R_1 -et a 3.1. tétel felhasználásával állítottuk elő, ezért $R_1 \subset G_h \subset G_k$.

Ekkor

$$R_1 + (F \wedge G_k) \subset G_k.$$

és p (2.5)-beli meghatározása miatt

$$\sum_{j \in R_1 + (F \wedge G_k)} a_j \cong \sum_{j \in G_k} a_j \cong b.$$

Ebből és a 4.1. lemmából következik, hogy

$$G = G_p - \{p\} \not\subseteq G_r,$$

tehát

$$q_p \cong q_r.$$

2. eset: $\bar{q} < q_p$.

Ekkor $G_p \subset G_h$, és mivel $R_1 \subset G_h$,

$$R_1 + (F \wedge G_p) = (R_1 \wedge G_h) + (F \wedge G_p) \supset (R_1 \wedge G_p) + (F \wedge G_p) = G_p.$$

Az utolsó egyenlőség azért teljesül, mert R_0 -t is a 3.1. tétel felhasználásával állítottuk elő, ezért

$$R_0 \wedge G_p = \emptyset.$$

Innen

$$\sum_{j \in R_1 + (F \wedge G_p)} a_j \cong \sum_{j \in G_p} a_j > b.$$

Ez viszont a 4.1. lemma miatt azt jelenti, hogy $G_r \subset G_p$, tehát

$$q_r \cong q_p.$$

3. eset: $\bar{q} = q_p$.

Ebben az esetben az állítás triviális, de a 4.1. tétel még pontosabb eredményt szolgáltat.

A következő lemmára és tételre az algoritmus kidolgozásakor lesz szükségünk.

4.3. LEMMA. Ha

$$q' = \max \{\bar{q}, q_p\}, \quad q'' = \min \{\bar{q}, q_p\},$$

és

$$\{j \in N: q' \cong q_j \cong q''\} \subset F,$$

akkor

$$q_r = q_p.$$

Bizonyítás. A 4.2. lemma bizonyításával azonos, csak az 1. esetben

$$R_1 + (F \wedge G_k) = G_k,$$

a 2. esetben

$$R_1 + (F \wedge G_p) = G_p.$$

A 3. eset a 4.1. tételből következik.

4.1. TÉTEL. Egy tetszőlegesen választott \bar{q} szorzó akkor és csak akkor egyezik meg a segítségével végzett méretcsökkentés után fennmaradó feladat legjobb q_r szorzójával, ha megegyezett a teljes feladat legjobb $q^* = q_p$ szorzójával, azaz

$$\bar{q} = q_r \langle = \rangle \bar{q} = q_p.$$

Bizonyítás. A tétel „csak akkor” része a 4.2. lemma következménye. Most a másik irányt bizonyítjuk.

Tegyük fel, hogy $\bar{q} = q_p$. Ebből és h definíciójából következik, hogy

$$\bar{G}_h \wedge G_p \subset E_p.$$

Most belátjuk, hogy

$$r \in \bar{G}_h \wedge G_p.$$

Mivel R_0 -t a 3.1. tétel és $\bar{q} = q_p$ felhasználásával állítottuk elő,

$$R_0 \wedge G_p = \emptyset.$$

Ezért

$$G_p \subset R_1 + (F \wedge G_p),$$

tehát

$$\sum_{j \in R_1 + (F \wedge G_p)} a_j \cong \sum_{j \in G_p} a_j > b,$$

ami azt jelenti, hogy

$$r \in G_p.$$

Másrészt a feltevés miatt

$$p \notin G_h,$$

tehát

$$\sum_{j \in R_1 + (F \wedge G_h)} a_j \cong \sum_{j \in G_h} a_j \leq b.$$

Így

$$r \in \bar{G}_h.$$

A fentiek ismeretében már leírhatjuk a q^* szorzót előállító algoritmus vázát. *PIVLOC (pivot-locating) algoritmus*

1. Becsüljük meg a q^* szorzót! Jelöljük a becslést \bar{q} -sal!

$$q_1 := +\infty, \quad q_2 := 0$$

2. Inicializáljuk a becslült optimum módszert, azaz futtassuk mindaddig, amíg annak 3. lépésében a fennmaradó feladatnak már lesz megengedett megoldás (Azaz amíg

$$\sum_{j \in R_1} a_j \leq b$$

nem teljesül.)

3. Keressük meg a fennmaradó feladat legjobb szorzóját!

(Legyen ez q_r .)

4. Ha

$$\bar{q} = q_r,$$

akkor

$$\bar{q} = q^* = q_p, \quad \text{STOP!}$$

5. Ha

$$\bar{q} > q_r,$$

akkor legyen

$$q_1 := \min \{\bar{q}, q_1\}, \quad q_2 := \max \{q_r, q_2\},$$

egyébként

$$q_1 := \min \{q_r, q_1\}, \quad q_2 := \max \{\bar{q}, q_2\},$$

és

$$\bar{q} := 0,5 * (q_1 + q_2),$$

menjünk a 2. lépésre.

A 3. pontbeli feladat legtöbbször különösen kis méretű lesz, mert a méret-csökkentésnél használt célfüggvénybecslés az optimálisnál is „jobb” lehet, sőt célszerű így megválasztani.

A 4.1. tétel biztosítja azt, hogy az algoritmus leáll a 4. lépésben, mihamarabb $\bar{q} = q^*$.

A 4.2. lemma alapján látjuk, hogy a \bar{q} becslések sorozata az iterációk számában exponenciálisan konvergál a q^* szorzóhoz.

Az algoritmust a következő két teszt segítségével tehetjük végessé:

1. *Teszt.* Ha egy iteráció után az 5. lépésben q_1 vagy q_2 valamelyikének értéke nem változott, (jelölje ezt q_0), és ez az érték eredetileg nem egy becslésből, hanem egy fennmaradó feladatból (q_r) származott, akkor legyen

$$\bar{q} := q_0$$

az algoritmus 5. pontjában szereplő érték helyett.

2. *Teszt.* Ha az 5. lépésben kapott (q_1, q_2) intervallum elég rövid, akkor vagy elfogadjuk a felezőpontját közelítően optimálisnak, vagy megvizsgáljuk, hogy az intervallum tartalmaz-e nem azonos q_j értékeket. Ha nem, akkor készen vagyunk, ha igen akkor folytathatjuk az eljárást.

A 2. teszt egyedül nyilvánvalóan biztosítja az algoritmus végességét.

Az 1. teszt előnye, hogy olcsó és előbbre hozhatja az eljárás befejezését. Ezt a 4.3. lemma és a következő segítségével láthatjuk be.

4.4. LEMMA. A PIVLOC algoritmus jelöléseit használva, ha annak 5. lépésében bármikor

$$q_r = q^*,$$

akkor az algoritmus további iterációiban

$$\text{vagy } q_1 = q^*,$$

$$\text{vagy } q_2 = q^*.$$

Bizonyítás. A 4.2. lemmából következik, hogy az eljárás folyamán

$$q_1 \cong q^* \cong q_2.$$

Ezért a feltétel legelső fennállása esetén ha

$$\bar{q} < q_r,$$

akkor

$$\min \{q_r, q_1\} = q_r = q^*,$$

ha

$$\bar{q} > q_r,$$

akkor

$$\max \{q_r, q_2\} = q_r = q^*.$$

Így az állítás már következik a q_1 és q_2 sorozatok monotonitásából és a $q_1 \cong q^* \cong q_2$ egyenlőségekből.

Ezek után bebizonyíthatjuk azt, hogy egy feltehetően nagyon gyakori esetben az 1. teszt elég ahhoz, hogy az algoritmus véges legyen. A probléma az, hogy csak utólag tudjuk meg fennállt-e ez az eset, vagy sem.

4.5. LEMMA. Ha a méretcsökkentéshez használt \bar{z} célfüggvénybecslés kisebb, mint a z_c folytonos optimális célfüggvényérték, akkor a legjobb $q^* = q_p$ szorzónak van olyan $V(q^*)$ nyílt környezete, hogy ebben tetszőleges \bar{q} szorzót választva, az azzal elvégzett méretcsökkentés után fennmaradó feladat legjobb q_r szorzója megegyezik q^* -gal, azaz, ha

$$\bar{z} < z_c,$$

akkor

$$\exists V(q^*): \forall \bar{q} \in V(q^*), q_r = q^*.$$

Bizonyítás. Mivel a feladat változóinak száma véges, ezért nyilvánvaló, hogy q^* -nak van olyan $V'(q^*)$ nyílt környezete, hogy

$$\{j \in N: q_j \in V'(q^*)\} = E_p.$$

Most megmutatjuk, hogy q^* -nak van olyan $V''(q^*)$ nyílt környezete is, amelyben tetszőleges \bar{q} szorzót választva

$$E_p \subset F.$$

Legyen

$$f_j(q) := M(q) - |c_j - qa_j|, \quad j = 1, \dots, n.$$

Az f_j függvény folytonos q -ban, így

$$\forall \varepsilon > 0, \exists \delta_j > 0: |q - q^*| < \delta_j \Rightarrow |f_j(q^*)| - |f_j(q)| \leq |f_j(q^*) - f_j(q)| < \varepsilon.$$

Könnyen látható, hogy

$$z_c = f_j(q^*)$$

minden $j \in E_p$ esetén.

Így

$$\varepsilon := z_c - \bar{z}$$

választása esetén minden

$$j \in E_p$$

indexre és minden olyan q -ra, amelyre

$$|q - q^*| < \delta_j,$$

azt kapjuk, hogy

$$\bar{z} < f_j(q).$$

A 3.1. tétel alapján tehát $j \in F$.

Legyen tehát

$$V''(q^*) := \bigcap_{j \in E_p} \{q: |q - q^*| < \delta_j\}!$$

Ekkor a

$$V(q^*) := V'(q^*) \wedge V''(q^*)$$

környezet a 4.3. lemma miatt kielégíti a követelményeket.

5. Alkalmazások

A bináris hátizsák feladat legkézenfekvőbb alkalmazása [3], egy hátizsák feltöltése különböző fajtájú, a_j súlyú és c_j kalóriaértékű konzervekkel úgy, hogy egy adott b súlyhatáron belül maximális kalóriamennyiség kerüljön bele. Egy másik közvetlen alkalmazás a beruházási probléma. Itt ismerjük az egyes beruházások megvalósításától várható c_j hasznot, valamint a beruházás a_j költségét. Ki kell választani n db beruházási lehetőség közül azokat, amelyek összességükben maximális hasznot eredményeznek, de megvalósításukkal nem lépünk túl egy adott beruházási keretet. ((1.1) feladat.)

A gyakorlatban a fenti közvetlen alkalmazásokra ritkán kerül sor, minthogy egyetlen feltétel általában kevés a valóság leírására. Ezért a hátizsák feladatot leginkább segédeszközként szokták használni bonyolultabb feladatok megoldása során. A továbbiakban ilyen alkalmazásokkal foglalkozunk.

5.1. DEFINÍCIÓ. A

$$(P_1) \quad \min \{f_1(x): x \in F_1\}$$

feladat relaxáltjának nevezzük a

$$(P_2) \quad \min \{f_2(x): x \in F_2\}$$

feladatot, ha

$$(i) \quad F_1 \subset F_2$$

$$(ii) \quad \forall x \in F_1, \quad f_1(x) \cong f_2(x).$$

Diszkrét programozási feladatok megoldása során gyakran szoktak alkalmazni relaxációt. A relaxált feladat legtöbbször lényegesen gyorsabban megoldható, mint az eredeti (csak ekkor van értelme a relaxálásnak), optimális célfüggvényértéke viszont alsó becslést szolgáltat az eredeti feladat optimumához. Ez a becslés jól használható például a korlátozás és szétválasztás típusú eljárásokban, de optimalitási kritériumként máshol is [21]. A következő eset a legkellemesebbek közé tartozik. Amennyiben

$$\forall x \in F_1 \quad f_1(x) = f_2(x),$$

akkor ha x^* a (P_2) feladat optimális megoldása és

$$x^* \in F_1,$$

akkor x^* a (P_1) feladatnak is optimális megoldása.

A hátizsák feladat alkalmazásának legegyszerűbb esete a többfeltételes 0–1 lineáris programozási feladatok relaxációja, ahol a feltételek tetszőleges nemnegatív lineáris kombinációja alkotja a hátizsák feladat egyetlen feltételét.

Most egy másik alkalmazást írunk le kissé részletesebben.

Tekintsük a következő általánosított hozzárendelési feladatot:

$$(5.2) \quad \begin{aligned} \min \quad & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \\ & \sum_{i \in I} x_{ij} = 1 \\ & j \in J = \{1, \dots, n\} \\ & \sum_{j \in J} a_{ij} x_{ij} \leq b_i \\ & i \in I = \{1, \dots, m\} \\ & x_{ij} \in \{0, 1\}, \quad \forall i \in I, j \in J. \end{aligned}$$

E modell például akkor használható, ha m különböző technológiai paraméterekkel rendelkező, de azonos rendeltetésű gép között kell szétosztanunk n munkát úgy, hogy minimális idő alatt végezzünk azon feltételek mellett, hogy egy munka csak egy gépre kerülhet és a gépek kapacitása korlátozott.

A következő relaxáció az ún. *Lagrange-féle relaxációk* közé tartozik (hasonló relaxációt alkalmaztunk a hátizsák feladat méretének csökkentéséhez):

$$(5.3) \quad \begin{aligned} \min \quad & \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} u_j \left(1 - \sum_{i \in I} x_{ij}\right) \\ & \sum_{j \in J} a_{ij} x_{ij} \leq b_i, \quad \forall i \in I \\ & x_{ij} \in \{0, 1\}, \quad \forall i \in I, j \in J, \end{aligned}$$

ahol $u \in R_+^n$.

Az (5.3) feladat olyan struktúrájú, hogy megoldása a következő n darab ($i = 1, \dots, n$) hátizsák feladat megoldását jelenti:

$$\begin{aligned} \min \quad & \sum_{j \in J} (c_{ij} - u_j) x_{ij} \\ & \sum_{j \in J} a_{ij} x_{ij} \leq b_i \\ & x_{ij} \in \{0, 1\}, \quad \forall j \in J. \end{aligned}$$

6. Számítógépes tapasztalatok

Az ismertett algoritmusokra SIMULA—67 nyelvű programok készültek, amelyeket az MTA CDC—3300-as gépen próbáltunk ki.

1. *Módszernek* a következő részekből álló programot nevezzük:

- rendezés,
- *Ingargiola és Korsh-féle méretcsökkentés*,
- a fennmaradó feladat megoldása GREENBERG és HEGERICH algoritmusával.

2. *Módszernek* a következő részekből álló programot nevezzük:

- a legjobb szorzó keresése,
- a becslült optimum módszer,
- a fennmaradó feladatok megoldása az 1. *módszerrel*.

A becslült optimum módszer kezdeti \bar{z}_1 célfüggvénybecslése a $z_c = M(q^*)$ folytonos optimum volt, amit a legjobb szorzót megadó eljárás szolgáltat. A becslés módosítást az 1. iterációs lépésben a következőképpen választottuk meg:

$$\varepsilon_i := \min\{c_j - a_j\} : j \in F\}$$

A táblázatban látható feladatméretek mindegyikéből 30 különböző feladatot generáltunk. Az együtthatók 1 és 999 között egyenletes eloszlású véletlen számok voltak. Minden feladatot három különböző jobb oldallal oldottunk meg. Ezek

$$0,25 * \sum_{j \in N} a_j,$$

$$0,5 * \sum_{j \in N} a_j,$$

$$0,75 * \sum_{j \in N} a_j.$$

A 2. *módszerben* a legjobb szorzó keresésére kipróbáltuk mind a QUICKSORT, mind a PIVLOC algoritmust. Ez utóbbinál a szorzó kezdeti becslését minden esetben 0,75-nek választottuk. A 2. *módszer* futási ideje mindkét szorzó kereséssel a változók számának lineáris függvényével volt közelíthető. (A regressziós egyenes egyenlete: $t = 3,58 \cdot n + 329,82$). Mivel a PIVLOC algoritmus nem bizonyult jobbnak a QUICKSORT eljárásnál, és ez utóbbi egyszerűbb is, ezért levonhatjuk azt a következtetést, hogy a legjobb szorzó keresését célszerűbb az utóbbival végrehajtani. A fenti ε_i választással a becslült optimum módszer utolsó becslése csaknem mindig beletalált a diszkrét optimumba. Ez az oka annak, hogy a legnagyobb fennmaradó feladatok mérete (maximális rendezett méret) rendkívül kicsi lett, és bár átlagosan 4—5 ilyen feladatot kellett megoldani (iterációs szám) a becslült optimum módszer alkalmazása során, a teljes megoldási idő a rendezési módszerének a 10-edére csökkent 1000 változó esetén.

Az alábbi táblázat a számítógépes kísérletek néhány jellemző eredményét foglalja össze. Az itt szereplő számok a méretenként megoldott 3×30 feladat alapján számolt átlagok. A szórások nem jelentősek, mégis megjegyezzük, hogy a futási idők a legnagyobb jobb oldalak esetén általában kisebbek voltak.

n	CPU időarányok		2. módszer átlagok		
	rendezés — 1. módszer	2. módszer — 1. módszer	CPU idő (msec)	maximális rendezett méret	iteráció szám
50	0,46	1,08	331,3	7,1	4,2
100	0,57	0,55	469,7	8,7	4,7
150	0,60	0,50	896,3	10,6	4,7
200	0,63	0,45	984,7	11,5	4,7
250	0,69	0,44	1480,3	12,5	4,8
500	0,74	0,23	2481,0	15,6	4,7
750	0,83	0,13	3098,0	16,5	4,4
1000	0,88	0,10	3649,3	17,0	4,2

IRODALOM

- [1] AHRENS, J. H. and FINKE, G., "Merging and sorting applied to the zero-one knapsack problem", *Opns. Res.* **23** (1975) 1099—1109.
- [2] BALAS, E. and ZEMEL, E., "Solving large zero-one knapsack problems", *Management Sciences Research Report* No. 408 *R Carnegie—Mellon University* (1977).
- [3] BÍRÓ, M., „A hátizsák feladat méretének csökkentése”, Szakdolgozat, ELTE TTK, Budapest, 1977.
- [4] BÍRÓ, M., "Efficient method applying incomplete ordering for solving the binary knapsack problem", in *Optimization Techniques* Ed. K. Iracki, M. Malanowski, S. Walukiewicz (Springer-Verlag, 1980) 160—169.
- [5] CABOT, A. V., "An enumeration algorithm for knapsack problems", *Opns. Res.* **18** (1970) 306—311.
- [6] DANTZIG, G. B., "Discrete-variable extremum problems", *Opns. Res.* **5** (1957) 266—277.
- [7] FAALAND, B., "Solution of the value-independent knapsack problem by partitioning", *Opns. Res.* **21** (1973) 332—337.
- [8] FAYARD, D. and PLATEAU, G., "Resolution of the 0—1 knapsack problem: Comparison of methods", *Math. Prog.* **8** (3) (1975) 272—307.
- [9] FAYARD, D. et PLATEAU, G., "Techniques de resolution du probleme du knapsack en variables bivalentes : partie 3", *Publications № 91 du Laboratoire de Calcul de l'Universite des Sciences et Techniques de Lille* (1977).
- [10] FORGÓ, F., *Nemkonvex és diszkrét programozás* (Közgazdasági és Jogi Könyvkiadó, Budapest, 1978).
- [11] GEOFFRION, A. M., "Lagrangean relaxation for integer programming", *Math. Prog. Study* **2** (1974) 82—114.
- [12] GILMORE, P. C. and GOMORY, R. E., "A linear programming approach to the cutting stock problem", *Opns. Res.* **9** (1961) 849—858.
- [13] GILMORE, P. C. and GOMORY, R. E., "A linear programming approach to the cutting stock problem II", *Opns. Res.* **11** (1963) 863—888.
- [14] GILMORE, P. C. and GOMORY, R. E., "Multi-stage cutting stock problems of two and more dimensions", *Opns. Res.* **13** (1965) 94—120.
- [15] GILMORE, P. C. and GOMORY, R. E., "The theory and computation of knapsack functions", *Opns. Res.* **14** (1966) 1045—1074.
- [16] GREENBERG, H. and HEGERICH, R. L., "A branch search algorithm for the knapsack problem", *Man. Sci.* **16** (5) (1970) 327—332.
- [17] HOROWITZ, E. and SAHNI, S., "Computing partitions with applications to the knapsack problem", *JACM* **21** (1974) 277—292.
- [18] INGARGIOLA, G. P. AND KORSH, J. F., „Reduction algorithmh for zero-one single knapsack problems", *Man. Shci.* **30** (4) (1973) 460—463.
- [19] JEROSLOW, R. G., "Trivial integer programs unsolvable by branch-and-bound", *Math. Prog.* **6** (1974) 105—109.

- [20] KOLESAR, P. J., "A branch and bound algorithm for the knapsack problem", *Man. Sci.* **13** (1967) 723—735.
- [21] KOVÁCS, L. B., *A diszkrét programozás kombinatorikus módszerei* (Bolyai János Matematikai Társulat, Budapest, 1969).
- [22] LAND, A. and DOIG, A., "An automatic method of solving discrete programming problems", *Econometrica* **28** (1960) 497—520.
- [23] LAUIERE, M., "An algorithm for the 0—1 knapsack problem", *Math. Prog.* **14** (1978) 1—10.
- [24] LEHMANN, E., *Testing Statistical Hypotheses* (John Wiley and Sons, New York, 1959).
- [25] MARTELLO, S. and TÓTH, P., "An upper bound for the zero-one knapsack problem and a branch and bound algorithm", *EJOR* **1** (1977) 169—175.
- [26] NAUSS, R. M., "An efficient algorithm for the 0—1 knapsack problem", *Man. Sci.* **23** (1) (1976) 27—31.
- [27] NEYMAN, J. and PEARSON, E. S., "On the problem of the most efficient tests of statistical hypotheses", *Philosophical Transactions of the Royal Society of London, Series A*, **231** (1933).
- [28] PRÉKOPÁ, A., *Lineáris programozás* (Bolyai János Matematikai Társulat, Budapest, 1968).
- [29] SALKIN, H. M. and DE KLUYVER, C. A., "The knapsack problem: A survey", *NRLQ* **22** (1975) 127—144.
- [30] SUHL, U., "An algorithm and efficient data structures for the binary knapsack problem", *EJOR* **2** (1978) 420—428.
- [31] VÍZVÁRI, B., „A Lagrange szorzók használata diszkrét programozási algoritmusokban”, *Alk. Mat. Lapok* **2** (1976) 413—425.
- [32] ZOLTNERS, A. A., "A direct descent binary knapsack algorithm", *JACM* **25** (2) (1978) 304—311.

(Beérkezett: 1982. augusztus 12.)

BÍRÓ MIKLÓS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1502 BUDAPEST XI., KENDE U. 13—17.

THE BINARY KNAPSACK PROBLEM

M. BÍRÓ

This paper presents a survey of the basic solution methods of the binary knapsack problem. The results of the author related to this topic are included. Some theorems are presented and proved in a more general but still applied form. If the knapsack problem lends itself for a simpler proof, that one is also presented. After some possibilities of application we give account of computational experiences.

EGY MÁTRIXEGYENLET MEGOLDÁSÁRÓL ÉS AZ EREDMÉNYEK ALKALMAZÁSÁRÓL VALÓSZÍNŰSÉGI ELOSZLÁSFÜGGVÉNYEK LINEÁRIS KOMBINÁCIÓINAK ELMÉLETÉBEN

GYIRES BÉLA

Debrecen

A dolgozat két fejezetre oszlik. Az elsőben egy mátrixegyenletet vizsgálunk megoldhatóságának szempontjából. Az eredményeket az 1.1. tétel tartalmazza. Erre vezethető vissza az 1.2. tétel igazolása, amely az ismert *Sherman—Morrison-formula* (1.9) kiterjesztése. — A második fejezetben az 1.2. tételre és [2] dolgozatra építve a 2.1. tételben szükséges és elégséges feltételt adunk arra, hogy adott valószínűségi eloszlásfüggvény adott valószínűségi eloszlásfüggvények pozitív együtthatókkal történő lineáris kombinációjával legyen előállítható.

1. Egy mátrixegyenlet megoldásáról

Amint az közismert, az ún. *Sherman—Morrison-formulának* ([1]) nagy szerepe van bizonyos mátrixok inverzének kiszámításában. Ebben a fejezetben egy mátrixegyenlet megoldásával foglalkozunk. A kapott eredmények speciális eseteként igazolunk egy azonosságot, amely a *Sherman—Morrison-formula* kiterjesztésének tekinthető.

Jelölje $C(m \times n)$ azoknak az m sorból és n oszlopból álló mátrixoknak a halmazát, amelyeknek elemei komplex számok. $A^* \in C(n \times m)$ jelöli az $A \in C(m \times n)$ mátrix transzponáltját és adj $B \in C(n \times n)$ a $B \in C(n \times n)$ mátrix adjungáltját. E és 0 az egységmátrixot, illetve a zéró mátrixot jelöli.

E fejezet főcélja a következő tétel igazolása:

1.1. TÉTEL. Legyen $A \in C(n, n)$. Ekkor az

$$(1.1) \quad \text{adj } A \text{ Det } (A + X) - \text{adj } (A + X) \text{ Det } A = \text{adj } AX \text{ adj } A$$

mátrixegyenletnek az $X=0$ triviális megoldáson kívül összes megoldása a következő:

a) Ha A szinguláris és $\text{adj } A=0$, akkor minden $X \in C(n, n)$ kielégíti az (1.1) egyenletet.

b) Legyen A szinguláris és $\text{adj } A \neq 0$. Ebben az esetben $X \in C(n, n)$ akkor és csak akkor elégíti ki az (1.1) egyenletet, ha eleget tesz a

$$(1.2) \quad \text{Det } (A + X) = b^* X a$$

egyenletnek, ahol $\text{adj } A = ab^*$ és $a, b \in C(n, 1)$.

c) Legyen A és $A + X$ reguláris. Ekkor (1.1) összes megoldása a következő:

(I) $X = AN$, ahol $N \in C(n, n)$ tetszés szerinti 2 indexű nilpotens mátrix.

$$(II) \quad X = \left(\exp \frac{2j\pi}{k-1} i - 1 \right) A P_k \quad (j=1, \dots, k-2; k=3, \dots, n).$$

ahol $P_k \in C(n, n)$ olyan idempotens mátrix, amelynek rangja k .

(III) $X = (y-1)AP_1$, ahol $y \neq 0, 1$, különben tetszés szerinti komplex szám lehet.

d) Ha A reguláris és $A+X$ szinguláris, akkor (1.1) egyenlet összes megoldását $X=uv^*$ adja meg, ahol $u, v \in C(n, 1)$ a $v^*A^{-1}u + 1 = 0$ feltételnek tesz eleget.

Bizonyítás. Mivel $X=0$ triviálisan elégíti ki (1.1) egyenletet, a következőkben feltesszük, hogy $X \neq 0$.

Ad a) Az állítás triviális.

Ad b) Ebben az esetben (1.1)

$$(1.3) \quad \text{adj } A \text{ Det } (A+X) = \text{adj } AX \text{ adj } A$$

alakra redukálódik. Mivel A szinguláris, ezért $\text{adj } A = ab^*$, ahol $a, b \in C(n, 1)$. (1.3) alapján

$$ab^* \text{ Det } (A+X) = ab^*b^*Xa,$$

ahonnan $ab^* \neq 0$ miatt arra jutunk, hogy (1.2) feltétel szükséges (1.3) megoldhatóságához. Ha (1.2) feltételből indulunk ki, (1.3) formulára jutunk vissza, azaz (1.2) egyben elegendő is (1.3) megoldhatóságához.

Ad c) és d) $Y=A+X$ jelöléssel (1.1) egyenlet az

$$(1.4) \quad \text{adj } A \text{ Det } Y - \text{adj } Y \text{ Det } A = \text{adj } A(Y-A) \text{ adj } A$$

alakra hozható. Szorozzuk meg balról (1.4) egyenletet A mátrixszal, vegyük figyelembe, hogy $A \text{ adj } A = E \text{ Det } A$, azután a kapott egyenletet osszuk el A determinánsával. Így azt kapjuk, hogy

$$(1.5) \quad A \text{ adj } Y + Y \text{ adj } A = E (\text{Det } A + \text{Det } Y).$$

Reguláris A esetén (1.1) és (1.5) egyenletek ekvivalensek.

Hogy célunk szempontjából (1.5) kifejezésnél alkalmasabbat találjunk, írjuk (1.5) egyenletet az

$$(Y-X) \text{ adj } Y + (A+X) \text{ adj } A = E (\text{Det } A + \text{Det } Y)$$

alakba. Felhasználva az

$$Y \text{ adj } Y = E \text{ Det } Y, \quad A \text{ adj } A = E \text{ Det } A$$

azonosságokat, arra jutunk, hogy

$$X[\text{adj } (A+X) - \text{adj } A] = 0,$$

illetve az

$$\text{adj } (A+X) = \text{adj } [A(E+A^{-1}X)] = \text{adj } (E+A^{-1}X) \text{ adj } A$$

azonosság alapján, hogy

$$(1.6) \quad X \text{ adj } (E+A^{-1}X) = X.$$

Mivel A reguláris, (1.6) egyenlet (1.5) egyenlettel, így (1.1) egyenlettel ekvivalens. Felhasználva az

$$(E+A^{-1}X) \text{ adj } (E+A^{-1}X) = E \text{ Det } (E+A^{-1}X)$$

azonosságot, (1.6) alapján az

$$X(E+A^{-1}X) = X \text{ Det } (E+A^{-1}X)$$

kifejezésre jutunk, ami pedig az

$$(1.7) \quad A^{-1}X[\text{Det}(E + A^{-1}X) - 1] = (A^{-1}X)^2$$

egyenlettel azonos. Viszont (1.7) akkor és csak akkor ekvivalens (1.6), így (1.1) egyenlettel, ha $A + X$ reguláris. Ha $A + X$ szinguláris, (1.7) szükséges feltétel ahhoz, hogy X (1.1) egyenletnek megoldása legyen.

Ad c) Ha $y = \text{Det}(E + A^{-1}X) = 1$, akkor (1.7) alapján $A^{-1}X = N$ tetszés szerinti 2 indexű nilpotens mátrix. Ha viszont $A^{-1}N$ 2 indexű nilpotens, akkor $E + N = \exp N$ alapján $\text{Det}(E + N) = \exp(\text{Sp } N) = 1$. Így (I) állításunkat igazoltuk.

Ha $y \neq 0, 1$, akkor (1.7) alapján $A^{-1}X = (y-1)P$, ahol $P \in C(n, n)$ idempotens. Innen

$$E + A^{-1}X = (y-1)P + E.$$

Mivel P idempotens, sajátértékei az 1 és 0 számok, továbbá hasonlósági transzformációval diagonális alakra hozható. Legyen $P = P_k$. Ekkor utolsó egyenletünkben $y = y^k$ és így $k \geq 3$ esetben $y \neq 0, 1$ feltétel miatt y 1-től különböző $(k-1)$ -edik egységgyök.

Ha $k=1$, akkor az $y^{k-1}=1$ egyenletet bármely $y \neq 0$ komplex szám kielégíti. Így $X = (y-1)AP_1$, ahol y tetszés szerinti, a 0 és 1 számoktól különböző komplex szám.

Ha még figyelembe vesszük, hogy $y \neq 0$ esetében

$$\text{Det}(E + (y-1)P_k) = y^k \neq 1,$$

ha

$$y = \exp \frac{2j\pi}{k-1} i \quad (j = 1, \dots, k-2, k \geq 3),$$

és $k=1$ esetben, ha $y \neq 1$, akkor (II) és (III) állításunk teljes bizonyításához jutunk.

Ad d) Ebben az esetben (1.7) alapján ahhoz, hogy X megoldása legyen az (1.1) egyenletnek, szükséges az $X = -AP$ feltétel teljesülése, ahol P idempotens mátrix. Hogy meghatározhassuk, ezek közül melyek azok, amelyek valóban kielégítik az (1.1) egyenletet, helyettesítsük be X eme értékét (1.1) egyenletünkbe. Rövid számolás után arra jutunk, hogy P mátrixnak ki kell elégítenie a $P = \text{adj}(E - P)$ feltételt. Mivel $A + X$ szinguláris, $E - P$ is az. Tehát $P = cb^*$, $b, c \in C(n, 1)$, amelynek 1 sajátértéke, azaz kell, hogy $b^*c = 1$ legyen. Legyen most $C = -A^{-1}a$, akkor $X = ab^*$, ahol $b^*A^{-1}a + 1 = 0$.

Ezzel az 1.1 tételt igazoltuk.

Az 1.1 tétel alapján könnyen igazolhatjuk a következő azonosságot:

1.2. TÉTEL. Ha $A \in C(n, n)$ és $u, v \in C(n, 1)$, és ha A szinguláris, $\text{adj } A = 0$, vagy ha A szinguláris, $\text{adj } A \neq 0$, vagy ha A és $A + uv^*$ reguláris, akkor

$$(1.8) \quad \text{adj } A \text{ Det}(A + uv^*) - \text{adj}(A + uv^*) \text{ Det } A = \text{adj } A uv^* \text{ adj } A.$$

Ha A reguláris és $A + uv^*$ szinguláris, akkor és csak akkor igaz (1.8), ha még $b^*A^{-1}a + 1 = 0$.

Bizonyítás. Csak azt kell kimutatni, hogy $X = uv^*$ kielégíti az 1.1 tétel a), b) és c) eseteiben megfogalmazott feltételeket.

Az $uv^* = 0$, továbbá a szinguláris A és $\text{adj } A = 0$ eset triviális.

Ha A szinguláris és $\text{adj } A = ab^* \neq 0$, akkor

$$\text{Det}(A + uv^*) = v^* ab^* u = b^*(uv^*)a,$$

azaz $X = uv^*$ kielégíti az (1.2) feltételt.

Legyen A és $A + uv^*$ reguláris. Ekkor

$$\text{Det}(E + A^{-1}uv^*) = 1 + v^* A^{-1}u,$$

így

$$A^{-1}uv^*[\text{Det}(E + A^{-1}uv^*) - 1] = v^* A^{-1}u \cdot A^{-1}uv^*,$$

azaz $X = uv^*$ az (1.7) feltételt is kielégíti. Ezzel az 1.2. tételt igazoltuk.

Ha A és $A + uv^*$ regulárisak, (1.8) azonosságot egyszerű átalakítással

$$(1.9) \quad (A + uv^*)^{-1} = A^{-1} - \frac{1}{1 + v^* A^{-1}u} A^{-1}uv^* A^{-1}$$

alakra hozhatjuk, és ez a *Sherman—Morrison formula*.

E tétel általánosításaként az 1.1. tétel alapján adódik a következő állítás:

1.1. KÖVETKEZMÉNY. Reguláris A , $A + X \in C(n, n)$ esetén

$$(A + X)^{-1} = A^{-1} - \frac{1}{\text{Det}(E + A^{-1}X)} A^{-1}XA^{-1}$$

előállítás akkor és csak akkor teljesedik, ha X az 1.1. tétel c) esetében felsorolt (I), (II) és (III) mátrixok közül bármelyik.

2. A valószínűségi eloszlásfüggvények szuperpozíciója

Jelölje V a valószínűségi eloszlásfüggvények (VEF-ek) halmazát. Legyen $S_n \subset C(n \times 1)$ azoknak a vektoroknak a halmaza, amelyeknek komponensei pozitív számok és összegük 1.

Azt mondjuk, hogy $F \in V$ szuperponálható a $G_j \in V$ ($j = 1, \dots, n$) VEF-ekkel, ha $F = \alpha_1 G_1 + \dots + \alpha_n G_n$, ahol $\alpha = (\alpha_j) \in S_n$.

Azt, hogy az elméleti vizsgálatokban és az alkalmazásokban milyen fontos szerepe van a VEF-ek szuperponálhatósági kérdésének, mutatja MEDGYESSY PÁLnak erről a tárgyköréről írt két monográfiája ([3], [4]). Szerző [2] dolgozatában e két monográfiában található tárgyalattól merőben különböző módszerekkel foglalkozik VEF-ek szuperponálhatóságával. [2] dolgozat egyik főeredménye szükséges és elegendő feltételek megadása annak eldöntésére, vajon adott VEF adott VEF-ekkel szuperponálható-e, vagy sem. Ebben a fejezetben ezt a tételt fogalmazzuk át az 1.2. tétel segítségével numerikus számítások szempontjából előnyösebb alakba.

E tétel ismertetéséhez bizonyos előkészültre van szükség.

Legyen (R, B) valamely mérhető tér, ahol R a valós számok halmaza, B e halmaz részhalmazainak *Borel-féle σ -algebrája*. Jelölje $\omega(F)$ a B σ -algebrán értelmezett azt a mértéket, amelyet $F \in V$ generál. Azt mondjuk, hogy $\omega(F)$ mérték abszolút folytonos a B halmazon értelmezett σ -véges λ mértékre, ha $\omega(F)(A) = 0$ mindazokra a B σ -algebrához tartozó A halmazokra, amelyekre $\lambda(A) = 0$.

A Radon—Nikodym-tétel értelmében $\omega(F)$ akkor és csak akkor abszolút folytonos a σ -véges λ mértékre, ha van olyan Borel-mérhető f függvény, hogy bármely $A \in \mathcal{B}$ esetén fennálljon a

$$\omega(F)(A) = \int_A f \lambda(dx)$$

egyenlőség. Ekkor f a $\omega(F)$ mértéknek a λ mértékre vonatkozó sűrűségfüggvénye, vagy másként a $\omega(F)$ mértéknek a λ mértékre vonatkozó Radon—Nikodym-deriváltja:

$$f = \frac{\omega(F)dx}{\lambda(dx)}.$$

Ha a következőkben nem is hangsúlyozzuk, ismételten felhasználjuk azt a tényt, hogy amennyiben $\{\omega(F_j)\}$, $F_j \in V$, a B σ -algebrán értelmezett véges, vagy megszámlálható végtelen sok σ -véges mérték, akkor létezik a B halmazon olyan σ -véges λ mérték, amelyre valamennyi $\omega(F_j)$ mérték abszolút folytonos.

A következőkben az $\int_{-\infty}^{\infty}$ helyett a \int jelet használjuk.

Legyen $F, G_j \in V$ ($j=1, 2$). Legyenek $\omega(F)$, $\omega(G_j)$ ($j=1, 2$) abszolút folytonosak a σ -véges λ mértékre, és legyenek f, g_j ($j=1, 2$) a megfelelő sűrűségfüggvények. A következőkben fontos szerepet játszik a következő funkcionál:

$$(2.1) \quad (G_1, G_2)_F = \int \frac{g_1 g_2}{f} \lambda(dx).$$

Ez az integrál egyértelműen meghatározott és független λ megválasztásától. Ha $G_1 = G_2 = G$, amikor egyben $g_1 = g_2 = g$, akkor a

$$(G, G)_F = \int \frac{g^2}{f} \lambda(dx)$$

mennyiséget a G VEF-nek az F VEF-re vonatkozó diszkrepanciájának, (2.1) mennyiséget pedig G_1, G_2 VEF-nek az F VEF-re vonatkozó együttes diszkrepanciájának nevezzük. Kimutatható, ([2], Theorem 2.1.), hogy $(G, G)_F \geq 1$, egyenlőség akkor és csak akkor, ha $G = F$. Továbbá az is, hogy $(G_1, G_2)_F > 0$ és ha $(G_j, G_j)_F$ ($j=1, 2$) végesek, akkor $(G_1, G_2)_F$ is az.

Jelölje a következőkben $H(F)$ azoknak a VEF-eknek a halmazát, amelyeknek az F VEF-re vonatkozó diszkrepanciája véges.

Legyen $F \in V$, $G_j \in H(F)$ ($j=1, \dots, n$).

$$\Gamma = \Gamma(G_j | F(j=1, \dots, n)) = ((G_j, G_k)_F)_{j,k=1}^n$$

mátrixot a G_j ($j=1, \dots, n$) VEF-eknek az F VEF-re vonatkozó Gram mátrixának nevezzük. Nyilván Γ szimmetrikus és belátható ([2], Theorem 3.1.), hogy pozitív szemidefinit.

Jelölje $e \in C(n \times 1)$ azt a vektort, amelynek minden komponense 1.

Igaz a következő állítás. Ha VEF-ek lineáris függetlenségét (függsőségét) a szokott módon értelmezzük, akkor a $G_j \in H(F)$ ($j=1, \dots, n$) VEF-ek akkor és csak

akkor lineárisan függetlenek, ha $\Gamma = \Gamma(G_j | F(j=1, \dots, n))$ mátrix pozitív definit ([2], Theorem 3.3.).

Most már az a tétel, amelynek számítástechnikai szempontból előnyösebb alakra hozását e fejezet céljaként jelöltük meg, a következő ([2], Theorem 4.9.):

Az $F \in V$ akkor és csak akkor szuperponálható a lineárisan független $G_j \in H(F)$ ($j=1, \dots, n$) VEF-ekkel, ha $e^* \Gamma^{-1} e = 1$, és ha $\text{adj } \Gamma e = (a_j)$ komponensei pozitívak, ahol Γ a G_j ($j=1, \dots, n$) VEF-eknek az F VEF-re vonatkozó Gram mátrixa. Ebben az esetben

$$(2.2) \quad F = \frac{1}{a} \sum_{j=1}^m a_j G_j, \quad a = e^* \text{adj } \Gamma e.$$

E tétel alapján igazoljuk a következő állítást:

2.1. TÉTEL. Legyen $F \in V$. Legyenek a $G_j \in H(F)$ ($j=1, \dots, n$) VEF-ek lineárisan függetlenek és jelölje Γ ezeknek a VEF-eknek az F VEF-re vonatkozó Gram mátrixát. F akkor és csak akkor szuperponálható a G_j ($j=1, \dots, n$) VEF-ekkel, ha $\Gamma - ee^*$ mátrix szinguláris és ha $b = (b_j)$ komponensei pozitívak, ahol b vektort $\text{adj } (\Gamma - ee^*) = bb^*$ azonosság határozza meg. Ebben az esetben

$$(2.3) \quad F = \sum_{k=1}^n \frac{b_k}{b_1 + \dots + b_n} G_k.$$

Számítástechnikai szempontból (2.3) előállítás (2.2) előállítással szemben előnyösebb. Ugyanis (2.2) az $\text{adj } \Gamma$ kiszámítását, tehát $\frac{1}{2} n(n+1)$ számú $(n-1)$ -edrendű aldetermináns kiszámítását teszi szükségessé. (2.3) előállítás viszont n számú $(n-1)$ -edrendű aldeterminánsét.

2.1. tétel bizonyítása. Mivel $\Gamma - ee^*$ szinguláris, Γ reguláris,

$$(2.4) \quad \text{Det } \Gamma = e^* \text{adj } \Gamma e > 0,$$

azaz $e^* \Gamma^{-1} e = 1$ és $\text{Det } (\Gamma - ee^*) = 1$ feltételek ekvivalensek. 1.2. tételt alkalmazva

$$(2.5) \quad \text{Det } \Gamma \text{adj } (\Gamma - ee^*) = \text{adj } \Gamma ee^* \text{adj } \Gamma,$$

ahonnan (2.4) segítségével azt kapjuk, hogy

$$(2.6) \quad e^* \text{adj } (\Gamma - ee^*) e = e^* \text{adj } \Gamma e.$$

Felhasználva az

$$\text{adj } \Gamma e = (a_j), \quad \text{adj } (\Gamma - ee^*) = bb^*$$

jelöléseket,

$$(2.7) \quad \begin{aligned} e^* \text{adj } (\Gamma - ee^*) e &= (b_1 + \dots + b_n)^2, \\ e^* \text{adj } \Gamma e &= a_1 + \dots + a_n = a. \end{aligned}$$

Ha (2.5) azonosságból kapott

$$\text{Det } \Gamma b_j b_k = a_j a_k$$

összefüggésben j -re összegezzünk és felhasználjuk (2.6) egyenlőségünket, az $a_k = (b_1 + \dots + b_n) b_k$, illetve az

$$\frac{a_k}{e^* \text{adj } \Gamma e} = \frac{(b_1 + \dots + b_n) b_k}{(b_1 + \dots + b_n)^2}$$

relációra jutunk. Figyelembe véve (2.7) azonosságokat,

$$\frac{a_k}{a} = \frac{b_k}{b_1 + \dots + b_n} \quad (k = 1, \dots, n),$$

ami már 2.1. tétel teljes igazolását jelenti.

IRODALOM

- [1] DWYER, P. S., and WAUGH, F. V., "On errors in matrix inversion", *J. Amer. Statist. Assoc.* **48** (1953) 289—319.
- [2] GYIRES, B., "Contribution to the theory of linear combinations of the probability distribution functions", *Studia Sci. Math. Hungar.* **16** (1983) Megjelenés alatt.
- [3] MEDGYESSY, P., *Decompositions and Superpositions of Distribution Functions* (Akadémiai Kiadó, Budapest, 1961).
- [4] MEDGYESSY, P., *Decompositions and Superpositions of Density Functions and Discrete Distributions* (Akadémiai Kiadó, Budapest, 1977).

(Beérkezett: 1982. december 14.)

GYIRES BÉLA

KÖSSUTH LAJOS TUDOMÁNYEGYETEM MATEMATIKAI INTÉZETE
4010 DEBRECEN, PF. 12.

ON THE SOLUTION OF CERTAIN MATRIX EQUATION, AND AN APPLICATION OF THE RESULTS IN THE THEORY OF LINEAR COMBINATIONS OF PROBABILITY DISTRIBUTION FUNCTIONS

B. GYIRES

The paper has two parts. In the first one it will be given necessary and sufficient conditions to obtain the solutions of the matrix equation (1.1). On the basis of this theorem we show that matrix equation (1.1) is satisfied by $X = uv^*$, where u and v are arbitrary column vectors, and v^* denotes the transpose of v . The identity (1.8) is an extension of the well-known formula (1.9) of *Sherman and Morrison*.

Denote by (R, B) the measurable space with *Borel σ -field* B generated by the set R of the real numbers. Denote by $\omega(F)$ the measure on B generated by the probability distribution function F . Let F, G_j ($j=1, \dots, n$) be probability distribution functions, and let the measures $\omega(F), \omega(G_j)$ ($j=1, \dots, n$), on B be dominated by the σ -finite measure λ on B , and let f, g_j ($j=1, \dots, n$) the corresponding density functions. We introduce the quantities

$$(G_j, G_k)_F = \int_{-\infty}^{\infty} \frac{g_j g_k}{f} \lambda(dx) \quad (j, k = 1, \dots, n)$$

and assume that $(G_j, G_j)_F$ ($j=1, \dots, n$) are finite. Let us introduce the matrix

$$\Gamma = ((G_j, G_k)_F)_{j,k=1}^n$$

Using identity (1.8), we give a proof of the following theorem in the second part of the paper on the basis of theorem 4.9. of [2]. Let $H(F) = \{G \in V | (G, G)_F < \infty\}$.

If $G_j \in H(F)$ ($j=1, \dots, n$) are linearly independent, then representation $F = \sum_{j=1}^n \alpha_j G_j$ holds with $\alpha_j > 0$ ($j=1, \dots, n$), $\sum_{j=1}^n \alpha_j = 1$ if and only if $\Gamma - ee^*$ is singular, and the components of vector b are positive numbers, where $\text{adj}(\Gamma - ee^*) = bb^*$, and e is the column vector with components 1. In this case

$$\alpha_j = \frac{b_j}{b_1 + \dots + b_n}, \quad (j = 1, \dots, n).$$

A SZINAJ-BILIÁRD POTENCIÁLTÉRBE. HOROCIKLUSKONSTRUKCIÓ

VETIER ANDRÁS

Budapest

A statisztikus fizikának a való világhoz közelebb álló dinamikus rendszerei közül elsőnek a *Szinaj-biliárdról* sikerült bebizonyítani az ergodicitást. (A *Szinaj-biliárd* fogalmát lásd [1]-ben, a bizonyítást pedig [2]-ben.) A bizonyítás négy fő részből áll:

- a) „horociklusok” konstrukciója,
- b) az „abszolút folytonosság” igazolása,
- c) a „Hopf-gondolatmenet” alkalmazása,
- d) az „Osznovnaja-lemma” bizonyítása.

[1]-ben, egy sokkal egyszerűbb modellel keresztül bemutattam, hogy mit jelenetnek az a)–d) pontokban szereplő fogalmak.

Ebben a cikkben általánosítom a *Szinaj-biliárdot*. Megengedem, hogy a biliárdasztalon egy potenciálfüggvény gradienseként fellépő erőter hasson. Szemléletesen ez azt jelenti, hogy a biliárdasztal nem sima, hanem a potenciálfüggvénynek megfelelően „hepe-hupás”. Dolgozatomban a horociklusok konstrukcióját végzem el erre a „hepe-hupás” biliárdra.

1. A biliárd fogalma

Tekintsünk egy kompakt, irányítható, perem nélküli, kétszer folytonosan differenciálható felületet. (Felületen kétdimenziós *Riemann-sokaságot* értünk.) Ilyen felület például a háromdimenziós euklideszi tér valamelyik gömbjének felülete, vagy az euklideszi sík valamelyik négyzetéből „a szemközti oldalak összeragasztásával” keletkező tórusz.

Ütközőnek hívjuk a felület olyan nyílt, összefüggő részalmazait, melyeknek határa kívülről nézve pozitív görbületű, kétszer folytonosan differenciálható görbe. Vegyünk véges sok ütközőt a felületen. *Biliárdasztalnak* hívjuk, és Q -val jelöljük a felületnek azt a részalmazát, mely a véges sok ütköző elhagyásával keletkezik a felületből. Ha a felület speciálisan a tórusz, akkor a biliárdasztalt *Szinaj-féle biliárdasztalnak* hívjuk. Az ütközők határainak egyesítése alkotja a *biliárdasztal szélét*, amit ∂Q -val jelölünk.

Képzeljük el, hogy egy pontszerű golyó mozog a biliárdasztalon úgy, hogy geodetikus vonalak mentén egységnyi sebességgel halad, amikor pedig a biliárdasztal széléhez ér, a rugalmas ütközés szabályai szerint visszaverődik az ütközőről, és az ennek megfelelő irányú geodetikus vonal mentén megy tovább, s í. t.

A golyó *állapotát* a helye és a sebessége határozza meg. Tehát egy x állapot egy q helyet és egy v irányt együtt jelent: $x=(q, v)$. Az $x=(q, v)$ állapot azonosítható a q pontban a v irányba mutató érintő-egységvektorral. *Fázistérnek* nevezzük, és M -mel jelöljük az összes lehetséges állapotok halmazát. Abban a pillanatban, amikor a golyó a biliárdasztal szélén ütközik, a golyó q -val jelölt helye egyértelműen definiált, de a sebessége nem. Ugyanis ha az ütköző falához v_- sebes

séggel érkezik, akkor az ütköző falától azzal a v_+ sebességgel verődik vissza, mely a v_- vektornak az ütköző q pontjában vett érintőjére való tükrösképe. Tekintve, hogy az ütközés időtartama nulla, a (q, v_-) és a (q, v_+) állapotokat azonosnak tekintjük.

A felület *Riemann-metrikája területmértéket* indukál a biliárdasztalon, amit μ_Q -val jelölünk. A biliárdasztal valamely q pontjában tekintett összes lehetséges irányok halmaza, az E^q halmaz kézenfekvő módon azonosítható a q pontbeli érintősík egységsugarú körvonalával. A körvonalon tekintett ívhosszmérték ennél az azonosításnál egy σ^q mértéknek felel meg E^q -n. Tekintve, hogy $M = \bigcup_{q \in Q} E^q$, a μ_Q területmértékből — mint vetületeloszlásból — és a σ^q ($q \in Q$) mértékekből — mint feltételes eloszlásokból — kapunk egy μ mértéket M -en.

A golyó mozgása egy $\{S^t\}_{t \in \mathbb{R}^1}$ dinamikát értelmez a fázistéren: ha a golyó most az $x \in M$ állapotban van, akkor $S^t x$ azt az állapotot jelenti, amiben a golyó t idő múlva lesz. Ha $t < 0$, akkor $S^t x$ a golyónak $|t|$ idővel ezelőtti állapotát jelenti. Ha tehát a golyó mostani állapota $x \in M$, akkor az $S^t x$ ($t \in \mathbb{R}^1$) állapotok a golyó múltbeli és jövőbeli állapotait, azaz a golyó *trajektóriáját* jelentik.

Mint ismeretes [4]-ből, a μ mérték invariáns az $\{S^t\}_{t \in \mathbb{R}^1}$ dinamikára. Ezért egy $(M, \{S^t\}_{t \in \mathbb{R}^1}, \mu)$ dinamikus rendszert kapunk, amit a biliárdasztalon tekintett *biliárdnak* nevezünk. Ha például a felület a tórusz, akkor speciális esetként a *Szinaj-biliárd*hoz jutunk. Ha pedig a felület *Gauss-gömbülete* nem azonosan nulla, akkor *görbült asztalon* tekintett biliárdról beszélünk.

Az 6. szakaszban definiálni fogjuk a potenciáltérben vett *Szinaj-biliárdot*, és ugyanott be fogjuk látni, hogy a potenciáltérben vett *Szinaj-biliárd* ergodikus szempontból ekvivalens egy görbült asztalon tekintett biliárddal. Ezért a 2—5. szakaszban csak a most definiált, „potenciáltértől mentes” biliárddal fogunk foglalkozni. (A 6. szakasz érthető a 2—5. szakaszok elolvasása nélkül is.)

Ismeretes (lásd például [1] vagy [2]), hogy az ergodicitás bizonyításánál milyen hasznos szerepet töltenek be a szűkülő és a táguló horociklusok. Ilyen horociklusokat fogunk definiálni a 2. szakaszban, a 3—5. szakaszban pedig megkonstruáljuk őket.

2. A horociklusok fogalma és az egzisztenciátétel kimondása

A fázistéren egy q *Riemann-metrikát* értelmezhetünk a biliárdasztalon vett Q_Q *Riemann-metrikából*. Egy görbét a fázistérben úgy kapunk, hogy veszünk egy g görbét a biliárdasztalon, és ennek minden egyes q pontjához egy $v(q)$ egységvektort illesztünk. Jelöljük $d\alpha$ -val a $v(q+dq)$ és $v(q)$ vektoroknak a g menti párhuzamos eltolásnál adódó szögét. A dq -nak megfelelő M -beli dx „elemi görbedarab” ívhossza legyen:

$$\varrho(dx) := \sqrt{(\varrho_Q(dq))^2 + (d\alpha)^2}.$$

Ha a $v(q)$ vektorok mind merőlegesek g -re, akkor nyilván $d\alpha = \kappa(q) \cdot \varrho(dq)$, ahol $\kappa(q)$ a g görbének a q pontbeli görbülete. Ezért a g -hez ily módon hozzárendelt \tilde{g} görbére

$$(2.1) \quad \tilde{g} \text{ ívhossza} = \int_g \sqrt{1 + (\kappa(q))^2} \varrho_Q(dq).$$

Vegyük egy x elemet a fázistérből. $H^{(sz)}(x)$ -et az x elemhez tartozó *szűkülő horociklusnak* nevezzük, ha $H^{(sz)}(x)$ folytonosan differenciálható görbe a fázistérben, az x elemet a belsejében tartalmazza, és minden $y \in H^{(sz)}(x)$ -re $S'y$ -nak $S'x$ -től való távolsága $t \rightarrow \infty$ esetén 0-hoz tart. A *táguló horociklus* definíciója ettől annyiban különbözik, hogy $t \rightarrow -\infty$ mellett követeljük meg, hogy $S'y$ és $S'x$ távolsága 0-hoz tartson.

Mint [1]-ben látható a „Bolyai-tóruszokon” vett geodetikus folyam fázisterének minden eleméhez végtelen hosszú szűkülő és táguló horociklus tartozik. Azt is megmutattam [1]-ben, hogy a szűkülő és a táguló horociklusok a *Bolyai-sík* geometriájából, nevezetesen az összesimuló geodetikus-seregekből egyszerűen adódnak.

Azonban már a *Szinaj-biliárdnál* is módosul kissé a helyzet, mert a fázistérnek csak majdnem minden eleméhez tartozik szűkülő és táguló horociklus, és ezeknek a hossza is csak véges (lásd [2]).

Egy későbbi példa mutatni fogja, hogy görbült asztalon tekintett biliárd esetén az ütközők pusztá léte nem elegendő ahhoz, hogy hasonlókat állíthassunk. Ezért feltételezzük, hogy a véges sok ütköző úgy helyezkedik el a felületen, hogy

1. bármely két ütköző távolsága pozitív, és
2. a biliárdgolyó ütközései közötti szabad úthossz felülről korlátos, azaz létezik olyan véges τ_{\max} érték, hogy a biliárdasztalon húzható geodetikus vonalak hossza kisebb vagy egyenlő, mint τ_{\max} . Jelöljük $k(q)$ -val a biliárdasztal szélének a görbületét ($q \in \partial Q$), és legyen $k_{\min} := \min_{q \in \partial Q} k(q) > 0$. Jelöljük $K(q)$ -val a biliárdasztal *Gauss-görbületét* ($q \in Q$), és legyen $K_{\max} := \max_{q \in Q} K(q)$.

Egzisztencia-tétel. Ha

$$(2.2) \quad K_{\max} \leq 0,$$

vagy

$$(2.3) \quad K_{\max} > 0, \quad \sqrt{K_{\max}} \tau_{\max} < \frac{\pi}{2} \quad \text{és} \quad \sqrt{K_{\max}} \operatorname{tg}(\sqrt{K_{\max}} \cdot \tau_{\max}) < k_{\min},$$

akkor a biliárd fázisterének majdnem minden eleméhez tartozik szűkülő és táguló horociklus.

A *Szinaj-biliárdban* a horociklusok léte az ütközők szóró hatásának a következménye (lásd [2]). Ha $K_{\max} \leq 0$, akkor a görbült asztal nem csökkenti ezt a szóró hatást, ezért minden további feltevés nélkül léteznek horociklusok. Ha azonban $K_{\max} > 0$, akkor ott, ahol a *Gauss-görbület* pozitív, összehúzó hatás érvényesül. Ezért egyáltalán nem nyilvánvaló, hogy ilyenkor is léteznek horociklusok. Ha viszont a (2.3) feltétel teljesül, vagyis a biliárdasztal *Gauss-görbülete* és a szabad úthossz elég kicsi az ütközők görbületéhez képest, akkor az ütközők szóró hatása legyőzi a biliárdasztal összehúzó hatását, és ezért ilyenkor — a *Szinaj-biliárdhoz* hasonlóan — léteznek horociklusok. Most egy példával megmutatjuk, hogy görbült asztalon tekintett biliárd esetén az ütközők pusztá léte nem elegendő ahhoz, hogy a fázistér majdnem minden eleméhez tartozzon szűkülő és táguló horociklus. Ha a gömbfelületen az ütközőket — mondjuk — az „északi sarkkörön belül” helyezzük el, akkor azok a trajektóriák, melyek az „északi sarkkör alatti főkörökre tekerednek fel”, együttesen a fázistér pozitív mértékű részét teszik ki. Nyilvánvaló, hogy azokhoz az elemekhez, melyek ilyen trajektóriákon vannak, nem tartozik sem szűkülő, sem táguló horociklus (1. ábra).



1. ábra

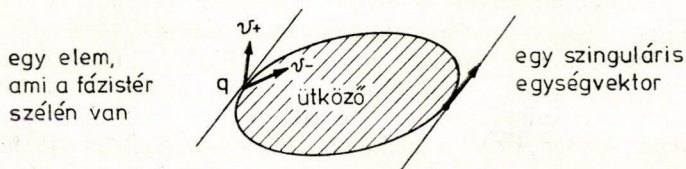
3. Fésűk és sugárnyalábok

A dolgozat hátra levő része a 6. szakasz kivételével az egzisztencia-tétel bizonyítása. Ezért a 3., 4. és 5. szakaszban, anélkül, hogy ezt később újra hangsúlyoznánk, mindvégig feltételezzük, hogy a (2.2) vagy a (2.3) feltétel teljesül. Magukat ezeket az egyenlőtlenségeket csak a 3.1. tétel bizonyításában használjuk fel, de ez a tétel annyira fontos szerepet tölt be a későbbiekben, hogy eltekintünk annak vizsgálatától, hogy mely definíciók és tételek lennének kimondhatók akkor is, ha ezek az egyenlőtlenségek nem teljesülnének.

Mint korábban leszögeztük, a fázistér olyan x elemekből áll, melyek a biliárdasztalnak egy q pontját és a q pontban vett valamilyen v irányt együttesen jelentenek: $x=(q, v)$. A v irány a biliárdasztal q pontjához tartozó érintősíkjának egy egységvektorával azonosítható, ezért a fázistér elemeit *egységvektoroknak* is fogjuk nevezni. A q pontot az x *egységvektor helyének* hívjuk. Ha q a biliárdasztal szélén van, akkor — mint korábban megállapodtunk — v egy irány-párt, az x tehát egy egységvektor-párt jelent (2. ábra). Az ilyen elemek alkotják a fázistér szélét. A fázistér szélének vannak olyan $x=(q, v)$ elemei, melyekre a $v=(v_-, v_+)$ iránypár tagjai egybeesnek. Ilyenkor a $v_-=v_+$ irány az ütköző érintőjének iránya. Az ilyen elemeket *szinguláris egységvektoroknak* hívjuk (2. ábra). Annak megfelelően, ahogy a biliárdasztal széle a véges sok ütköző ∂Q_i határának az egyesítése: $\partial Q = \bigcup_i \partial Q_i$, a fázistér széle is véges sok ∂M_i komponens egyesítése: $\partial M = \bigcup_i \partial M_i$.

Azt a $\pi: M \rightarrow Q$ leképezést, amely egy egységvektorhoz hozzárendeli a vektor helyét a biliárdasztalon, *természetes vetítésnek* nevezzük. Tehát ha $x=(q, v) \in M$, akkor $\pi(x)=q \in Q$.

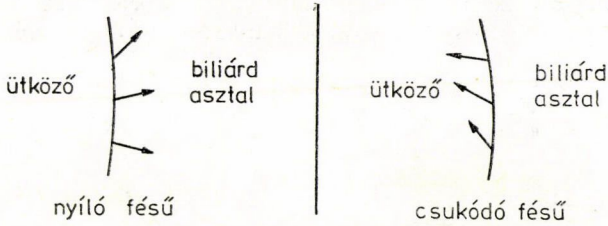
Tekintsünk egy $x \in M$ elemet. Az x elem trajektóriájának π szerinti vetületét a biliárdasztalon az x elem *pályájának* nevezzük. A pályának azt a paraméterezé-



2. ábra

sét, melynél a $\pi(S^1x)$ ponthoz a t paraméterérték tartozik ($t \in R^1$), a pálya *természetes paraméterezésének* nevezzük. Tehát egy x elem pályája a természetes paraméterezéssel az a görbe a biliárdasztalon, mely azt mutatja, hogy mikor hol van a golyó feltéve, hogy a golyó most az x állapotban van.

Szemeljünk ki egy $x \in M$ elemet, és legyen $\tau := \min \{t > 0; S^t x \in \partial M\}$. τ tehát a legkisebb olyan pozitív időpont, amikor az x állapotból indított golyó a biliárdasztal szélén van. τ -t az x elem jövőjében az *első ütközési időpontnak* nevezzük.



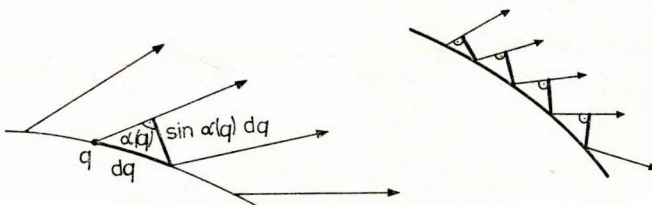
(Az ábrákon az egységvektor-pároknak csak az egyik tagját tüntettük fel.)

3. ábra

Tekintsünk egy ütközőt, és az ütköző határán egy $D(\alpha)$ szakaszt. Minden $q \in D(\alpha)$ ponthoz rendeljünk hozzá egy $\alpha(q) \in \pi^{-1}(q)$ elemet. Mivel $q \in \partial Q$, $\alpha(q)$ egy egységvektor-párt jelent a q pontban. A kapott $\alpha: D(\alpha) \rightarrow \partial M$ leképezést *fésűnek* nevezzük. Az ütköző határán lehetséges két irányítás valamelyikének irányában tekintsük $D(\alpha)$ -n az ívhossz-paramétert. Az $\alpha(q)$ elemnek az ütköző irányított határával bezárt szögét jelöljük $\underline{\alpha}(q)$ -val ($q \in D(\alpha)$). A kapott $\underline{\alpha}: D(\alpha) \rightarrow [0, \pi]$ függvényt a fésű *nyomának* nevezzük. Ha az α fésű nyoma folytonosan differenciálható és monoton csökkenő (vagy növekedő), akkor az α fésűt *nyíló* (vagy *csukódó*) fésűnek nevezzük (3. ábra). Ha például az α fésű nyoma konstans, akkor az α fésű *nyíló* is és *csukódó* is.

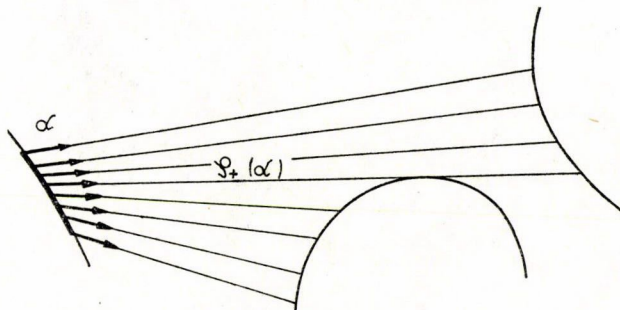
Az α fésű nyomának definíciójában szerepe van annak, hogy az ütköző határát hogyan irányítjuk. Könnyű meggyőződni róla, hogy egy α fésű *nyíló* vagy *csukódó* mivolta, nyomának az ívhossz-paraméter szerinti $\underline{\alpha}'(q)$ deriváltja, továbbá $\sin \underline{\alpha}(q)$ értéke nem függ az ütköző határának irányításától.

A $D(\alpha)$ szakaszt a fésű *értelmezési tartományának* nevezzük. $D(\alpha)$ hosszát a fésű *hosszának*, a hossz felét a fésű *sugarának* hívjuk. A $D(\alpha)$ szakasz felezőpontját a fésű *középpontjának* mondjuk. A $W(\alpha) = \int_{D(\alpha)} \sin \underline{\alpha}(q) dq$ értéket, ahol az integrálás az ívhossz szerint értendő, a fésű *szélességének* nevezzük (4. ábra). Látszik,



4. ábra

hogy ha $\alpha(q) = \frac{\pi}{2}$ ($q \in D(\alpha)$), akkor a fésű szélessége egyenlő a hosszával. Legyen $\alpha: D(\alpha) \rightarrow \partial M$ nyíló fésű. Mint már leszögeztük, $\alpha(q)$ egy $(v_-(q), v_+(q))$ egységvektor-párt jelent: $v_-(q)$ felel meg egy golyónak az ütközőhöz való becsapódási sebességének, $v_+(q)$ pedig a visszaverődési sebességének ($q \in D(\alpha)$). Minden egyes $q \in D(\alpha)$ ponthoz tekintsük az $\alpha(q)$ elem pályájának az első ütközési időpontig terjedő darabját, tehát húzzuk meg a q pontból $v_+(q)$ irányban induló geodetikus vonalat a biliárdasztal széléig. A kapott pályadarabok seregét az α fésűből induló *sugárnyaláb*nak nevezzük, és $\mathcal{S}_+(\alpha)$ -val jelöljük (5. ábra). A sugárnyaláb elemeit,

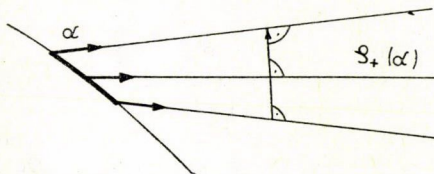


5. ábra

tehát a pályadarabokat, *sugaraknak* hívjuk. Világos, hogy az $\alpha(q)$ elem jövőjében az első ütközési időpont egyenlő a q pontból induló sugár hosszával ($q \in D(\alpha)$).

Az $\mathcal{S}_+(\alpha)$ -beli sugarak pontjainak halmazát $\mathcal{P}_+(\alpha)$ -val fogjuk jelölni. $\mathcal{P}_+(\alpha)$ a biliárdasztalnak az a részhalmaza, melyet az α fésűből induló sugárnyaláb *betakar*.

Ha α nyíló fésű, akkor $\mathcal{S}_+(\alpha)$ -nak van ortogonális görbeserege. Az ortogonális görbesereg elemeit irányítsuk úgy, hogy irányuk a sugarak irányából $+90^\circ$ -os elforgatással keletkezzen (6. ábra). Az ortogonális görbesereg valamely $q \in \mathcal{P}_+(\alpha)$



6. ábra

ponton átmenő elemének a q pontban vett előjeles görbületét a sugárnyaláb q pontbeli *szóródásának* nevezzük. $k(q)$ -val jelöltük az ütköző határának az előjel nélküli görbületét ($q \in \partial Q$). Egyszerű számolással adódik (lásd például [2], 2. §), hogy az α fésű értelmezési tartományának $q \in D(\alpha)$ pontjaiban az $\mathcal{S}_+(\alpha)$ sugárnyaláb szóródása

$$\frac{k(q) - \alpha'(q)}{\sin \alpha(q)}.$$

Ha a szóródást egy τ hosszúságú sugár mentén a t ívhosszparaméter függvénye-

ként tekintjük, akkor az így értelmezett $\kappa: [0, \tau] \rightarrow \mathbb{R}^1$ függvény eleget tesz az alábbi differenciálegyenletnek:

$$\kappa'(t) = -K(t) - \kappa^2(t),$$

ahol $K(t)$ -vel jelöltük a biliárdasztal *Gauss-görbületét* a sugár t paraméterű pontjában. (Ez a differenciálegyenlet levezethető például a [3] függelékében található ún. *Jacobi-egyenletből*.)

Ha a kiszemelt sugár a $q \in D(\alpha)$ pontból indul, akkor a fentiek szerint

$$\kappa(0) = \frac{k(q) - \alpha'(q)}{\sin \alpha(q)}.$$

Az α fésű nyoma monoton csökkenő, ezért $\alpha'(q) \leq 0$, s így $\kappa(0) \geq k_{\min} > 0$ ($q \in D(\alpha)$). Tehát egy nyíló fésűből induló sugárnyaláb „kezdőszóródása” legalább akkora, mint egy bizonyos pozitív konstans, és ezért a sugárnyaláb sugarai „széttartóan” indulnak ki a fésűből. Az alábbi tétel mutatja, hogy a sugárnyaláb „széttartó tulajdonsága” a következő ütközőig megőrződik.

3.1. TÉTEL. *Létezik olyan $\omega > 0$ konstans, mely csak a biliárdasztaltól függ, hogy $\kappa(t) \geq \omega$ ($0 \leq t \leq \tau$).*

A tétel szerint egy nyíló fésűből induló sugárnyaláb szóródása a sugárnyaláb által betakart halmazon mindenhol nagyobb vagy egyenlő, mint egy univerzális, pozitív konstans.

Bizonyítás. A $\kappa_{\min}(t)$ függvényt az alábbi differenciálegyenlettel definiáljuk:

$$\kappa'_{\min}(t) = -K_{\max} - \kappa_{\min}^2(t) \quad (0 \leq t \leq \tau_{\max}),$$

$$\kappa_{\min}(0) = k_{\min}.$$

Mivel $-K(t) - \kappa^2 \geq -K_{\max} - \kappa^2$ és $\kappa(0) \geq \kappa_{\min}(0)$, a megfelelő differenciálegyenletek megoldásaira is hasonló értelmű egyenlőtlenség áll fenn: $\kappa(t) \geq \kappa_{\min}(t)$ ($0 \leq t \leq \tau$). $K_{\max} > 0$ esetén

$$\kappa_{\min}(t) = \sqrt{K_{\max}} \operatorname{tg}(\varphi - \sqrt{K_{\max}} \cdot t),$$

ahol φ a kezdeti feltételből adódó $\sqrt{K_{\max}} \cdot \operatorname{tg} \varphi = k_{\min}$ egyenlet legkisebb pozitív megoldása. A (2.3) feltételből következik, hogy $\varphi > \sqrt{K_{\max}} \cdot \tau_{\max}$. Ebből és a tangens-függvény monotonitásából következik, hogy $\kappa_{\min}(\tau_{\max}) > 0$, és $0 \leq t \leq \tau_{\max}$ -ra

$$\kappa_{\min}(t) \geq \kappa_{\min}(\tau_{\max}) =: \omega.$$

$K_{\max} = 0$ esetén

$$\kappa_{\min}(t) = \frac{1}{t + \frac{1}{k_{\min}}} \geq \frac{1}{\tau_{\max} + \frac{1}{k_{\min}}} =: \omega.$$

$K_{\max} < 0$, $k_{\min} > \sqrt{-K_{\max}}$ esetén

$$\kappa_{\min}(t) = \sqrt{-K_{\max}} \operatorname{cth}(\varphi + \sqrt{-K_{\max}} \cdot t) \geq \sqrt{-K_{\max}} =: \omega.$$

$K_{\max} < 0$, $k_{\min} = \sqrt{-K_{\max}}$ esetén

$$\kappa_{\min}(t) = \sqrt{-K_{\max}} =: \omega.$$

$K_{\max} < 0$, $k_{\min} < \sqrt{-K_{\max}}$ esetén

$$\kappa_{\min}(t) = \sqrt{-K_{\max}} \operatorname{th}(\varphi + \sqrt{-K_{\max}} \cdot t) \cong \kappa_{\min}(0) = k_{\min} =: \omega.$$

Látható, hogy ω mindegyik esetben pozitív, csak a biliárdasztaltól függ, és $\kappa(t) \cong \omega$ ($0 \leq t \leq \tau$).

A következő három tétel felső becsléseket ad egy nyíló fésűből induló sugárnyaláb szóródására. Ezekre a becslésekre a későbbiekben lesz szükség. τ és $\kappa(t)$ ($0 \leq t \leq \tau$) ugyanazt jelenti továbbra is, mint az előzőekben. Legyen $K_{\min} := \min_{q \in Q} K(q)$ a biliárdasztal Gauss-görbületének a minimuma.

3.2. TÉTEL. $t \geq t_0$ -ra $\kappa(t) \leq \sqrt{|K_{\min}|} \cdot \operatorname{cth}(\sqrt{|K_{\min}|} \cdot t_0)$.

Bizonyítás. A $\kappa_{\max}(t)$ függvényt az alábbi differenciálegyenlettel értelmezzük:

$$\kappa'_{\max}(t) = |K_{\min}| - \kappa_{\max}^2(t) \quad (0 \leq t \leq \tau_{\max}),$$

$$\kappa_{\max}(0) = \infty.$$

A differenciálegyenlet megoldása $\kappa_{\max}(t) = \sqrt{|K_{\min}|} \cdot \operatorname{cth}(\sqrt{|K_{\min}|} \cdot t)$. Mivel $-K(t) - \kappa^2 \leq -K_{\min} - \kappa^2 \leq |K_{\min}| - \kappa^2$, és $\kappa(0) \leq \kappa_{\max}(0)$, a megfelelő differenciálegyenletek megoldásaira is hasonló értelmű egyenlőtlenség áll fenn: $\kappa(t) \leq \kappa_{\max}(t)$ ($0 \leq t \leq \tau$). Mivel a kotangenshiperbolikus-függvény $(0, \infty)$ -ben monoton csökkenő, kapjuk, hogy $t \geq t_0$ -ra $\kappa(t) \leq \kappa_{\max}(t) \leq \kappa_{\max}(t_0)$.

Legyen $\tau_{\min} > 0$ az ütközők közötti távolságok minimuma. Vezessük be a következő jelölést:

$$\Omega_0 := \sqrt{|K_{\min}|} \cdot \operatorname{cth}(\sqrt{|K_{\min}|} \cdot \tau_{\min}).$$

3.3. TÉTEL. $\kappa(\tau) \leq \Omega_0$.

A tétel szerint egy nyíló fésűből induló sugárnyaláb szóródása a sugarak végpontjaiban kisebb vagy egyenlő, mint egy univerzális, véges konstans.

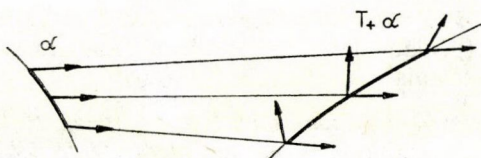
Bizonyítás. $\tau \geq \tau_{\min}$, ezért az előző tételből következik az egyenlőtlenség.

3.4. TÉTEL. $\kappa(t) \leq \max(\kappa(0); \sqrt{|K_{\min}|})$ ($0 \leq t \leq \tau$).

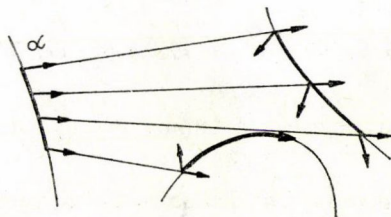
A tétel szerint egy nyíló fésűből induló sugárnyaláb szóródása a sugárnyaláb valamely sugara mentén sehol sem lehet nagyobb, mint a „kezdőszóródásnak” és egy univerzális konstansnak a maximuma.

Bizonyítás. Ha volna olyan $t \in [0, \tau]$, melyre $\kappa(t) > \kappa(0)$ és $\kappa(t) > \sqrt{|K_{\min}|}$, akkor volna olyan $t \in [0, t]$, melyre $\kappa'(t) > 0$ és $\kappa(t) > \sqrt{|K_{\min}|}$. Ilyen t viszont nem lehet, mert ha $\kappa(t) > \sqrt{|K_{\min}|}$, akkor $\kappa'(t) = -K(t) - \kappa^2(t) \leq -K_{\min} - \kappa^2(t) \leq \leq |K_{\min}| - \kappa^2(t) < 0$.

Egy nyíló α fésűből induló $\mathcal{S}_+(\alpha)$ sugárnyaláb sugarainak végpontjai egy vagy több ütköző határán egy vagy több szakaszt alkotnak. A sugarak végpontjaiban a sugarakhoz húzott érintő egységvektorok pedig egy vagy több fésűt definiálnak.



nem szakad szét



szétszakad

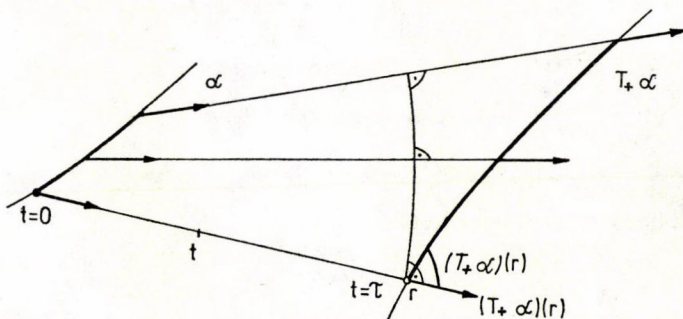
7. ábra

Ezeket a fésűket együttesen az α fésű T_+ -transzformáltjának nevezzük, és $T_+\alpha$ -val jelöljük. Ha az α fésű T_+ -transzformáltja csak egy fésűből áll, akkor azt mondjuk, hogy az α fésű *nem szakad szét* a T_+ -transzformációnál (7. ábra). Világos, hogy ha a nyíló α fésű *nem szakad szét* a T_+ -transzformációnál, akkor a $T_+\alpha$ fésű nyoma is folytonosan differenciálható. Egyszerű számolással adódik (lásd például [2], 2. §), hogy az $\mathcal{S}_+(\alpha)$ sugárnyaláb kiszemelt sugarának végpontjában a sugárnyaláb $\kappa(\tau)$ -val jelölt szóródása és a $T_+\alpha$ fésű nyoma között az alábbi kapcsolat áll fenn (8. ábra):

$$(T_+\alpha)'(r) = -k(r) - \kappa(\tau) \cdot \sin(T_+\alpha)(r) \quad (r \in D(T_+\alpha)).$$

Legyen $k_{\max} := \max_{q \in \partial Q} k(q)$, és $\Omega_1 := k_{\max} + \Omega_0$. Felhasználva a 3.1. és a 3.3. tételeket $(T_+\alpha)'(r)$ -re becslést kapunk:

$$(3.1) \quad -\Omega_1 \leq (T_+\alpha)'(r) \leq -k_{\min} \quad (r \in D(T_+\alpha)).$$



8. ábra

Ebből kiderül, hogy a $T_+\alpha$ fésű nyíló, és nyomának deriváltja univerzális konstansok közé esik. Ismeretes (például a [3] függelékében található ún. *Jacobi-egyenletből* levezethető) az alábbi lemma.

3.5. LEMMA. *Ha egy sugárnyaláb valamelyik sugarának t_1 , illetve t_2 paraméterű pontjában az ortogonális trajektóriák mentén mérve tekintjük egy másik, „közeli” sugártól való d_1 , illetve d_2 távolságokat (9. ábra), akkor*

$$\lim_{d_1, d_2 \rightarrow 0} \frac{d_2}{d_1} = \exp \int_{t_1}^{t_2} \kappa(t) dt,$$

ahol $\kappa(t)$ -vel a sugárnyaláb szóródását jelöltük a kiszemelt sugár t paraméterű pontjában.

A későbbiekben hasznos lesz, ha a lemma két nyilvánvaló következményét kihangsúlyozzuk.

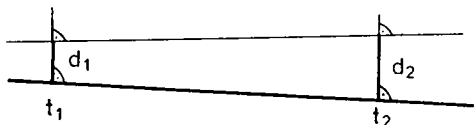
3.6. KÖVETKEZMÉNY. *Ha a sugárnyaláb szóródása a sugárnyaláb által betakart pontokban mindenhol legalább ω , akkor*

$$\lim_{d_1, d_2 \rightarrow 0} \frac{d_2}{d_1} \geq \exp(\omega \cdot (t_2 - t_1)),$$

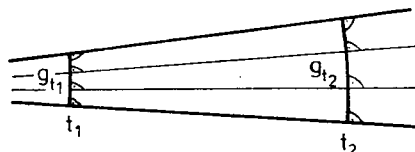
és így

$$\lim_{d_1, d_2 \rightarrow 0} \frac{d_2}{d_1} \geq 1 \quad (t_1 < t_2).$$

3.7. KÖVETKEZMÉNY. *Ha a sugárnyaláb szóródása a sugárnyaláb által betakart pontokban mindenhol legalább ω , akkor a sugárnyalábnak egy kiszemelt sugár t_1 paraméterű pontjához tartozó g_{t_1} , illetve t_2 paraméterű pontjához tartozó g_{t_2}*



9. ábra



10. ábra

ortogonális trajektóriáinak ívhosszaira (10. ábra) fennáll, hogy

$$g_{t_2} \text{ ívhossza} \geq g_{t_1} \text{ ívhossza} \cdot \exp(\omega \cdot (t_2 - t_1)),$$

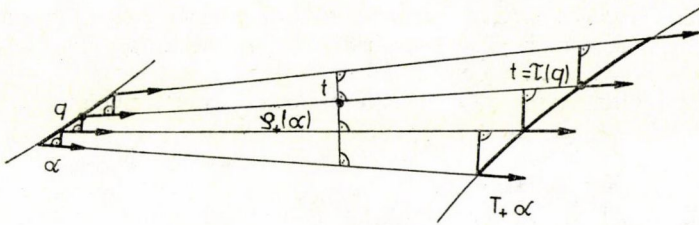
és így

$$g_{t_2} \text{ ívhossza} \geq g_{t_1} \text{ ívhossza} \quad (t_1 < t_2).$$

Vezessük be a következő jelölést:

$$A := \exp(\omega \cdot \tau_{\min}).$$

3.8. TÉTEL. *Ha a nyíló α fésű nem szakad szét a T_+ -transzformációnál, akkor $W(T_+\alpha) \geq A \cdot W(\alpha)$ (11. ábra).*



11. ábra

A tétel szerint egy nyíló fésű T_+ -transzformáltjának a szélessége nagyobb vagy egyenlő, mint a fésű szélességének egy 1-nél nagyobb konstansszorosa. Tehát egy nyíló fésű szélessége a T_+ -transzformációnál „exponenciálisan nő”.

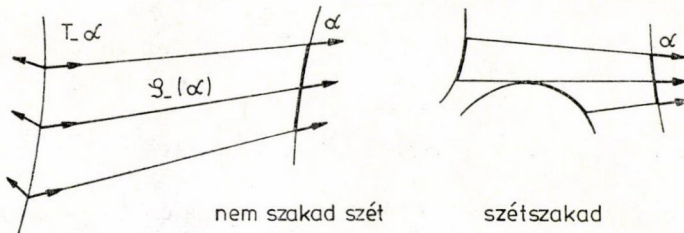
Bizonyítás. Legyen $\tau(q)$ a q pontból indított sugár hossza, $\kappa(q, t)$ pedig az $\mathcal{S}_+(\alpha)$ sugárnyalábnak a q pontból induló sugár t paraméterű pontjában vett szóródása ($0 \leq t \leq \tau(q)$, $q \in D(\alpha)$). Az előbb kimondott lemma alatt a sugarak menti megfeleltetésnél a $D(T_+\alpha)$ halmazon vett $\sin(T_+\alpha)(r) \cdot dr$ mértéknek a $D(\alpha)$

halmazon vett $\sin \alpha(q) \cdot dq$ mértékre vonatkozó sűrűségfüggvénye $\exp \int_0^{\tau(q)} \kappa(q, t) dt$

($q \in D(\alpha)$). Mivel $\kappa(q, t) \geq \omega$ és $\tau(q) \geq \tau_{\min}$ ($0 \leq t \leq \tau(q)$, $q \in D(\alpha)$), ebből következik,

$$\text{hogy } W(T_+\alpha) = \int_{D(\alpha)} \left(\exp \int_0^{\tau(q)} \kappa(q, t) dt \right) \cdot \sin \alpha(q) dq \geq \exp(\omega \cdot \tau_{\min}) \cdot \int_{D(\alpha)} \sin \alpha(q) dq = A \cdot W(\alpha).$$

Mindaz, amit nyíló fésűre elmondtunk, értelemszerűen átvittethető csukódó fésűre, csak nem a fésűből kiinduló, hanem a fésűbe beérkező sugarakkal kell mindent csinálni. Ezért nem részletezzük, csak egy-egy ábrával utalunk rá (12. ábra), hogy



12. ábra

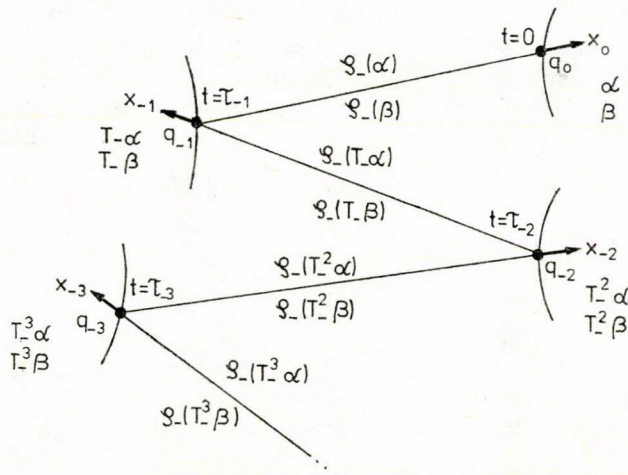
ha α csukódó fésű, akkor az α fésűbe beérkező $\mathcal{S}_-(\alpha)$ sugárnyalábnak, a sugárnyaláb által betakart $\mathcal{P}_-(\alpha)$ halmaznak, az α fésű T_- -transzformáltjának, illetve a T_- -transzformációnál szét nem szakadó fésűnek mi a definíciója.

¶ Nyilvánvaló, hogy csukódó fésűkre a T_- -transzformációval kapcsolatban teljesülnek azoknak az állításoknak az analógjai, melyek nyíló fésűkre igazak. Ezért a 3.1—3.8. tételek analógjait, melyeknek megfogalmazásától is eltekintünk, bizonyítottak vehetjük.

A most következő két tételt csukódó fésűkre a T_- -transzformációval kapcsolatban mondjuk ki, mert csak ilyen formában fogjuk később felhasználni őket.

Egy csukódó α fésű T_- -transzformáltja is csukódó. Ezért $T_- \alpha$ -nak is tekintjük a T_- -transzformációját, amit jelöljünk $T_-^2 \alpha$ -val. És így tovább kapjuk a $T_-^N \alpha$ fésűket ($N=1, 2, \dots$). Legyen $T_-^0 \alpha := \alpha$.

Tegyük fel, hogy az α és β fésűk csukódóak, és értelmezési tartományuknak egy közös q_0 elemén értékeik is egyenlőek. $\alpha(q_0) = \beta(q_0) := x_0$. Az x_0 elem múltjában az *utolsó, utolsó előtti, ... ütközési időpontok* definíciója kézenfekvő. Jelöljük ezeket $\tau_{-1}, \tau_{-2}, \dots$ -vel. Nyilván $\dots \tau_{-2} < \tau_{-1} < \tau_0 := 0$. Legyen $x_{-N} := S^{\tau_{-N}} x_0 \in D(T_-^N \alpha)$, $q_{-N} := \pi(x_{-N})$ ($N=1, 2, \dots$). Tekintsük az x elem pályáján a természetes ívhossz-paraméterezést, melynél a q_{-N} pont paramétere τ_{-N} ($N=1, 2, \dots$). Jelöljük $\kappa_{\mathcal{S}_-(T_-^N \alpha)}(t)$ -vel, illetve $\kappa_{\mathcal{S}_-(T_-^N \beta)}(t)$ -vel az $\mathcal{S}_-(T_-^N \alpha)$, illetve az $\mathcal{S}_-(T_-^N \beta)$ sugárnyaláb szóródását ennek a pályának a t paraméterű pontjában ($\tau_{-N-1} \leq t \leq \tau_{-N}$, $N=0, 1, \dots$) (13. ábra).



13. ábra

3.9. TÉTEL.

$$|\kappa_{\mathcal{S}_-(T_-^N \alpha)}(\tau_{-N}) - \kappa_{\mathcal{S}_-(T_-^N \beta)}(\tau_{-N})| \leq \frac{\Omega_0}{A^{2N-2}} \quad (N=1, 2, \dots).$$

Tehát az α és β csukódó fésűk T_-^N -transzformáltjaiba érkező sugárnyalábok szóródásának eltérése N növekedtével exponenciálisan csökken.

A tétel bizonyítását segéd-tételekre építjük.

3.10. SEGÉD-TÉTEL.

$$(3.2) \quad |\kappa_{\mathcal{S}_-(\alpha)}(\tau_{-1}) - \kappa_{\mathcal{S}_-(\beta)}(\tau_{-1})| \leq \Omega_0,$$

$$(3.3) \quad |\kappa_{\mathcal{S}_-(\alpha)}(\tau_{-1}) - \kappa_{\mathcal{S}_-(\beta)}(\tau_{-1})| \leq |\kappa_{\mathcal{S}_-(\alpha)}(0) - \kappa_{\mathcal{S}_-(\beta)}(0)| \cdot \frac{1}{A^2}.$$

Bizonyítás. (3.2) következik abból, hogy a 3.1. tétel analógja szerint $\kappa_{\mathcal{S}_-(\alpha)}(\tau_{-1})$ és $\kappa_{\mathcal{S}_-(\beta)}(\tau_{-1})$ is kisebb vagy egyenlő, mint $-\omega$, és a 3.3. tétel analógja szerint nagyobb vagy egyenlő, mint $-\Omega_0$.

(3.3) bizonyítása céljából tegyük fel, hogy $\kappa_{\mathcal{G}_{-(\alpha)}}(0) \equiv \kappa_{\mathcal{G}_{-(\beta)}}(0)$. Mivel $\kappa_{\mathcal{G}_{-(\alpha)}}(t)$ és $\kappa_{\mathcal{G}_{-(\beta)}}(t)$ ugyanannak a differenciálegyenletnek tesz eleget, a $\kappa_{\mathcal{G}_{-(\alpha)}}(0) \equiv \kappa_{\mathcal{G}_{-(\beta)}}(0)$ egyenlőtlenségből következik, hogy $\kappa_{\mathcal{G}_{-(\alpha)}}(t) \equiv \kappa_{\mathcal{G}_{-(\beta)}}(t)$ ($\tau_{-1} \leq t \leq 0$). Vegyük a

$$\begin{aligned}\kappa'_{\mathcal{G}_{-(\alpha)}}(t) &= -K(t) - \kappa_{\mathcal{G}_{-(\alpha)}}^2(t) \\ \kappa'_{\mathcal{G}_{-(\beta)}}(t) &= -K(t) - \kappa_{\mathcal{G}_{-(\beta)}}^2(t)\end{aligned} \quad (\tau_{-1} \leq t \leq 0)$$

differenciálegyenletek különbségét. Felhasználva, hogy a 3.1. tétel analógja szerint $\kappa_{\mathcal{G}_{-(\alpha)}}(t), \kappa_{\mathcal{G}_{-(\beta)}}(t) \leq -\omega$ ($\tau_{-1} \leq t \leq 0$), azt kapjuk, hogy

$$\begin{aligned}\kappa'_{\mathcal{G}_{-(\alpha)}}(t) - \kappa'_{\mathcal{G}_{-(\beta)}}(t) &= (-\kappa_{\mathcal{G}_{-(\alpha)}}(t) - \kappa_{\mathcal{G}_{-(\beta)}}(t)) \cdot (\kappa_{\mathcal{G}_{-(\alpha)}}(t) - \kappa_{\mathcal{G}_{-(\beta)}}(t)) \equiv \\ &\equiv 2\omega \cdot (\kappa_{\mathcal{G}_{-(\alpha)}}(t) - \kappa_{\mathcal{G}_{-(\beta)}}(t)) \quad (\tau_{-1} \leq t \leq 0).\end{aligned}$$

A $\Delta\kappa(t) := \kappa_{\mathcal{G}_{-(\alpha)}}(t) - \kappa_{\mathcal{G}_{-(\beta)}}(t)$ különbségre tehát fennáll, hogy

$$\Delta\kappa'(t) \equiv 2\omega \cdot \Delta\kappa(t) \quad (\tau_{-1} \leq t \leq 0).$$

Ha itt egyenlőség állna, akkor $\Delta\kappa(t)$ egyenlő volna $\Delta\kappa(0) \cdot \exp(2\omega t)$ -vel. Mivel egyenlőtlenség áll, $\Delta\kappa(t) \leq \Delta\kappa(0) \cdot \exp(2\omega t)$. Felhasználva, hogy $\tau_{-1} \leq -\tau_{\min}$ kiadódik a (3.3) képlet: $\Delta\kappa(\tau_{-1}) \leq \Delta\kappa(0) \cdot \exp(2\omega\tau_{-1}) \leq \Delta\kappa(0) \cdot \exp(-2\omega\tau_{\min}) = \Delta\kappa(0) \cdot \frac{1}{A^2}$.

3.11. SEGÉDTÉTEL.

$$\kappa_{\mathcal{G}_{-(T-\alpha)}}(\tau_{-1}) - \kappa_{\mathcal{G}_{-(T-\beta)}}(\tau_{-1}) = \kappa_{\mathcal{G}_{-(\alpha)}}(\tau_{-1}) - \kappa_{\mathcal{G}_{-(\beta)}}(\tau_{-1}).$$

Bizonyítás. „Az ütközőkről való visszaverődés szabályaiból” elemi eszközökkel adódik (lásd például [2], 2. §), hogy

$$\kappa_{\mathcal{G}_{-(\alpha)}}(\tau_{-1}) - \kappa_{\mathcal{G}_{-(T-\alpha)}}(\tau_{-1}) = \frac{2 \cdot k(q_{-1})}{\sin(T-\alpha)(q_{-1})}.$$

Ha ugyanezt β -val kapcsolatban is felírjuk, és a két egyenlőtlenséget egymásból kivonjuk, megkapjuk a bizonyítandó egyenlőséget.

3.12. SEGÉDTÉTEL.

$$(3.4) \quad |\kappa_{\mathcal{G}_{-(T-\alpha)}}(\tau_{-1}) - \kappa_{\mathcal{G}_{-(T-\beta)}}(\tau_{-1})| \leq \Omega_0,$$

$$(3.5) \quad |\kappa_{\mathcal{G}_{-(T-\alpha)}}(\tau_{-1}) - \kappa_{\mathcal{G}_{-(T-\beta)}}(\tau_{-1})| \leq |\kappa_{\mathcal{G}_{-(\alpha)}}(0) - \kappa_{\mathcal{G}_{-(\beta)}}(0)| \cdot \frac{1}{A^2}.$$

Bizonyítás. A 3.10. és a 3.11. segéd-tételekből közvetlenül adódnak ezek az egyenlőtlenségek.

3.13. SEGÉDTÉTEL.

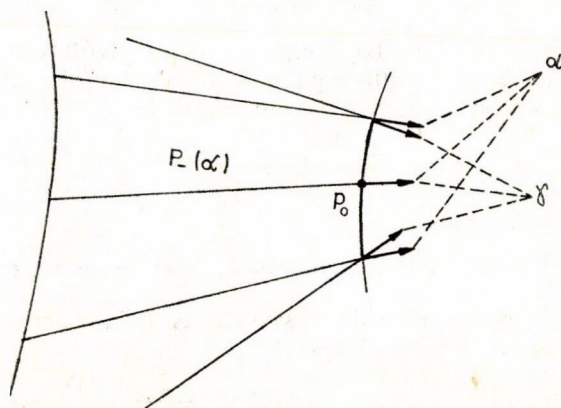
$$(3.6) \quad \begin{aligned} &|\kappa_{\mathcal{G}_{-(T_{-}^{n+1}\alpha)}}(\tau_{-(n+1)}) - \kappa_{\mathcal{G}_{-(T_{-}^{n+1}\beta)}}(\tau_{-(n+1)})| \leq \\ &|\kappa_{\mathcal{G}_{-(T_{-}^n\alpha)}}(\tau_{-n}) - \kappa_{\mathcal{G}_{-(T_{-}^n\beta)}}(\tau_{-n})| \cdot \frac{1}{A^2} \quad (n = 0, 1, \dots) \end{aligned}$$

Bizonyítás. Ez az egyenlőtlenség ugyanazt fejezi ki, mint a (3.5) egyenlőtlenség, csak α helyett $T^n \alpha$ -ra vonatkozik, és ennek megfelelően 0 helyett τ_{-n}, τ_{-1} helyett pedig $\tau_{-(n+1)}$ szerepel benne.

A 3.9. tétel bizonyítása. Írjuk fel a (3.6) egyenlőtlenséget $n=1, \dots, (N-1)$ -re. Ebből az $N-1$ darab egyenlőtlenségből és a (3.4) egyenlőtlenségből a tétel egyenlőtlensége adódik.

Ha α csukódó fésű, akkor korábbi jelöléseinknek megfelelően $\mathcal{P}_-(\alpha)$ jelöli a biliárdasztalnak azt a részhalmazát, melyet az $\mathcal{S}_-(\alpha)$ sugárnyaláb betakar.

3.14. TÉTEL. *Tegyük fel, hogy α konstans nyomú fésű, mely a T_- -transzformációnál nem szakad szét, γ pedig csukódó fésű ugyanazon az értelmezési tartományon. Tegyük fel még, hogy a közös értelmezési tartomány valamely p_0 elemén $\alpha(p_0) = \gamma(p_0)$. Ekkor $\mathcal{P}_-(\alpha) \subset \mathcal{P}_-(\gamma)$ (14. ábra).*



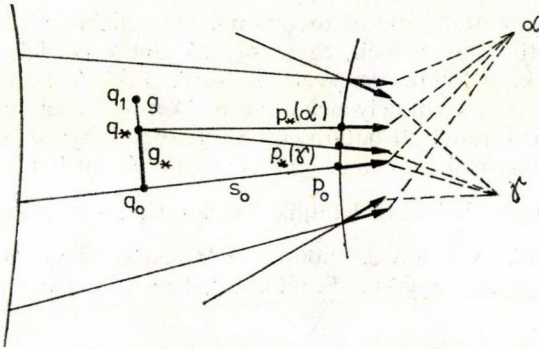
14. ábra

Bizonyítás. Indirekt bizonyítást végzünk: tegyük fel, hogy $\exists q_1 \in \mathcal{P}_-(\alpha) \setminus \mathcal{P}_-(\gamma)$. Az $\mathcal{S}_-(\alpha)$ sugárnyalábnak a p_0 pontba befutó sugarát, ami $\mathcal{S}_-(\gamma)$ -nak is eleme, jelöljük s_0 -lal. Tekintsük azt a legrövidebb $g \subset \mathcal{P}_-(\alpha)$ görbét, mely q_1 -et s_0 -lal összeköti. Ennek kezdőpontja $q_0 \in s_0$, végpontja q_1 . $g_* \subset g$ legyen az a maximális görbe, melynek kezdőpontja q_0 , és $g_* \subset \mathcal{P}_-(\gamma)$. g_* végpontja $q_* \in \mathcal{P}_-(\gamma)$. Indirekt feltevésünk miatt $q_* \neq q_1$. $s_*(\alpha) \in \mathcal{S}_-(\alpha)$, illetve $s_*(\gamma) \in \mathcal{S}_-(\gamma)$ legyenek azok a sugarak, melyek átmennek q_* -on; ezeknek a sugaraknak a végpontja $p_*(\alpha)$, illetve $p_*(\gamma)$ (15. ábra).

Mivel α nyoma konstans, a γ fésű pedig csukódó, a $p_*(\gamma)$ pont p_0 és $p_*(\alpha)$ közé esik. Ebből viszont következik, hogy q_* -nak egy kis környezete g -n $\mathcal{P}_-(\gamma)$ -ba esik, ami ellentmond g_* definíciójának.

A 3.14. tételből nyilvánvalóan adódik az alábbi

3.15. Következmény. A 3.14. tételben mondott feltételek mellett $D(T_- \alpha) \subset D(T_- \gamma)$.



15. ábra

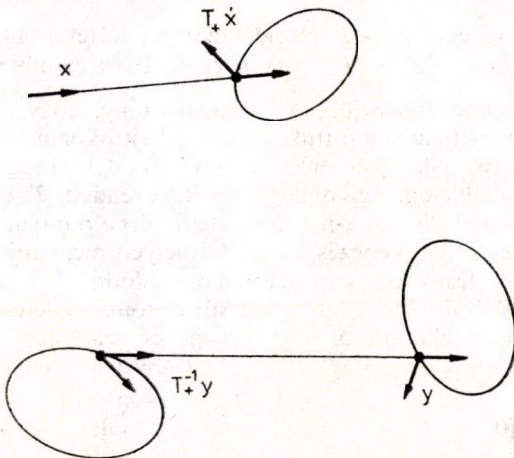
4. A deriváltfolyamat szingularitásai

Jelöljük $\tau(x)$ -szel az $x \in M$ elem jövőbeli első ütközési időpontját, tehát $\tau(x) := \min \{t > 0: S^t x \in \partial M\}$. A $T_+: M \rightarrow \partial M$ leképezés definíciója: $T_+ x := S^{\tau(x)} x$ ($x \in M$). A T_+ leképezésnek a ∂M -re való megszorítását az $\{S^t\}_{t \in \mathbb{R}^1}$ dinamika *deriváltfolyamatának* nevezzük, és ugyancsak T_+ -szal jelöljük. Világos, hogy a deriváltfolyamatnak van inverze, amit T_+^{-1} -gyel jelölünk, $T_+^{-1}: \partial M \rightarrow \partial M$ (16. ábra). Nyilvánvaló, hogy T_+^{-1} folytonos minden olyan y helyen, melyre $T_+^{-1}y$ nem szinguláris egységvektor. Ezért T_+^{-1} szakadási helyeit megkapjuk, ha a szinguláris egységvektorokra a T_+ leképezést alkalmazzuk.

Ha $X \subset \partial M$ és $0 < \Delta t \leq \tau_{\min}$, akkor $\mu(\{S^t x: x \in X, 0 < t \leq \Delta t\})$ értéke arányos Δt -vel, ezért a

$$v(X) := \frac{1}{\Delta t} \cdot \mu(\{S^t x: x \in X, 0 < t \leq \Delta t\}) \quad (X \subset \partial M)$$

képlettel egy v mértéket definiálhatunk ∂M -en. Ez a v nyilván invariáns T_+ -ra.

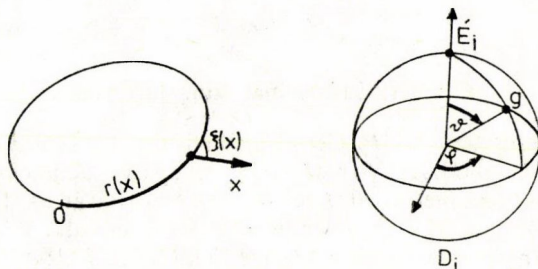


16. ábra

A derivált-folyamattal kapcsolatos számítások céljából minden ∂M_i komponensnek megfeleltetjük egy R^3 -beli, egységsugarú gömb G_i felületét. Ezt a következőképpen tesszük. Jelöljük \ker_i -vel az i -edik ütköző kerületének ívhosszát. Jelentse $r(x)$ az $x \in \partial M_i$ elem helyének az ütköző kerületén vett ívhossz paraméterét, $\xi(x)$ pedig az ütköző irányított határával bezárt szögét. Egy $x \in \partial M_i$ elemnek megfeleltetjük a G_i gömbfelületnek azt a $g = g(x)$ pontját, melynek polárkoordinátái:

$\varphi = \frac{r(x)}{\ker_i} \cdot 2\pi$, $\vartheta = \xi(x)$ (17. ábra). Jelöljük \vec{E}_i -vel a G_i gömb „északi sarkát”, D_i -vel

pedig a „déli sarkát”. Nyilvánvaló, hogy $x \in \partial M_i$ esetén $\xi(x) = 0$ vagy π akkor és csak akkor, ha $g(x) = \vec{E}_i$ vagy D_i . Tehát a ∂M_i -beli szinguláris egységvektorok két pontra képződnek.



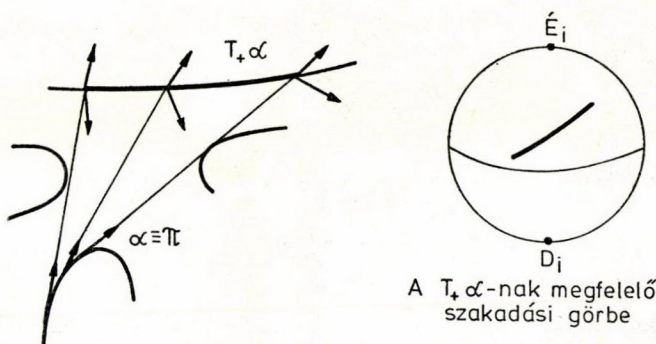
17. ábra

Eltekintve a szinguláris egységvektoroktól és az északi és a déli sarkoktól g kölcsönösen egyértelmű megfeleltetés. Ily módon a fázistér $\partial M = \bigcup_i \partial M_i$ határát véges sok G_i gömbfelülettel reprezentálhatjuk. Ennél a reprezentációnál a v mérték annak a mértéknek felel meg $\bigcup_i G_i$ -n, melynek a G_i -n vett megszorítása nem egyéb, mint a G_i -n vett felszínmérték $\frac{\ker_i}{2\pi}$ -szerese.

Mostantól kezdve egy $x \in \partial M$ elemet azonosnak tekintjük a neki megfelelő $g(x) \in \bigcup_i G_i$ ponttal, és a T_+ és T_+^{-1} leképezéseket is a gömbfelületek egyesítésén értelmezett leképezéseknek tekintjük. Az a zavaró tény, hogy a gömbök északi és déli sarkait több elemmel azonosítottuk, nem fog bajt okozni.

Világos, hogy egy nyíló fésű megfelel egy $\vartheta = \vartheta(\varphi)$ függvénynek, és ezáltal egy görbének a gömbfelületen. Egy nyíló α fésűhöz rendelt $T_+ \alpha$ fésűnek megfelelő görbe pontjait megkapjuk, ha az α -nak megfelelő görbe pontjaira a T_+ leképezést alkalmazzuk. Ezért a T_+^{-1} leképezés szakadási helyeit megkapjuk, ha az azonosan 0 vagy π nyomú α fésűk transzformáltjaiként adódó $T_+ \alpha$ fésűknek megfelelő görbéket — a *szakadási görbéket* — megkeressük a gömbfelületeken. A G_i gömbök északi és déli sarkait, valamint a T_+^{-1} leképezés szakadási görbéinek pontjait *szinguláris pontoknak* nevezzük. Egy $x \in M$ elem *sávjellemzőjét* így definiáljuk:

$$sj(x) := \sup \{c \in [0, 1] : T_+^n x \text{-nek a } \frac{c}{n^2} \text{ sugarú nyílt környezetében a gömbfelületen nincs } T_+^{-1} \text{-nek szinguláris pontja } (n=1, 2, \dots)\}.$$



18. ábra

A most következő 4.1. tételből és az 5. szakasz végén levő 5.9. tételből látható, hogy a sávjellemző hasznos fogalom lesz számunkra.

4.1. TÉTEL. A fázistér majdnem minden x elemére $s_j(x) > 0$.

Bizonyítás. Tudjuk, hogy egy nyíló α fésű $T_+\alpha$ transzformáltjának nyoma szigorúan monoton csökkenő. Ebből következik, hogy a $T_+\alpha$ fésűnek megfelelő görbe megadható $\varphi = \varphi(\vartheta)$ alakban. (3.1)-ből adódik, hogy $|(T_+\alpha)'| \cong k_{\min}$, amiből következik, hogy a $T_+\alpha$ -nak megfelelő $\varphi = \varphi(\vartheta)$ görbére:

$$\frac{d\varphi}{d\vartheta} = \frac{1}{\left| \frac{d\vartheta}{d\varphi} \right|} = \frac{\frac{2\pi}{\ker_i}}{|(T_+\alpha)'|} \cong \frac{2\pi}{\ker_i \cdot k_{\min}}.$$

Ezért ennek a görbének a gömbfelületen tekintett ívhossza véges:

$$\text{ív hossz} = \int_{\varphi_1}^{\varphi_2} \sqrt{1 + \left(\frac{d\vartheta}{d\varphi} \right)^2} \cdot \sin \vartheta d\varphi \cong \pi \cdot \sqrt{1 + \left(\frac{2\pi}{\ker_i \cdot k_{\min}} \right)^2}.$$

A szinguláris pontok halmazát B -vel, a B halmaz ε sugarú környezetét a gömbfelületeken B_ε -nal jelöljük.

Mivel a szinguláris pontok halmaza véges sok véges ívhosszú görbéből és véges sok pontból áll, létezik olyan, ε -tól nem függő C konstans, hogy

$$v(B_\varepsilon) \leq C \cdot \varepsilon \quad (\varepsilon > 0).$$

Felhasználva, hogy v invariáns T_+ -ra, kapjuk, hogy

$$\sum_n v(T_+^{-n} B_{\frac{1}{n^2}}) = \sum_n v(B_{\frac{1}{n^2}}) \leq C \cdot \sum_n \frac{1}{n^2} < \infty,$$

tehát a Borell—Cantelli-lemma miatt v.m.m. $x \in \partial M$ -hez $\exists n_0(x) < \infty$, hogy

$$T_+^n x \notin B_{\frac{1}{n^2}} \quad (n > n_0(x)).$$

Másrészt nyilván v -m.m. $x \in \partial M$ -re $T_+^n x \notin B$ ($n = 1, 2, \dots$), és így v -m.m. $x \in \partial M$ -hez $\exists c(x) \in (0, 1]$, hogy

$$T_+^n x \notin B_{\frac{c(x)}{n^2}} \quad (n = 1, \dots, n_0(x)).$$

E két tényt összevetve adódik, hogy v -m.m. $x \in \partial M$ -hez $\exists c(x) \in (0, 1]$, hogy

$$T_+^n x \notin B_{\frac{c(x)}{n^2}} \quad (n = 1, 2, \dots).$$

v és μ kapcsolatából kiadódik, hogy μ -m.m. $x \in M$ -hez is $\exists c(x) \in (0, 1]$, hogy

$$T_+^n x \notin B_{\frac{c(x)}{n^2}} \quad (n = 1, 2, \dots).$$

A sávjellemző definíciójából következik az alábbi két tétel.

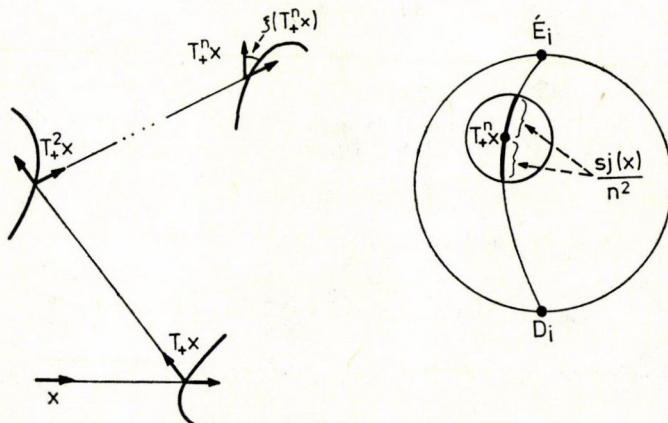
4.2. TÉTEL.

$$\frac{s_j(x)}{n^2} \leq \xi(T_+^n x) \leq \pi - \frac{s_j(x)}{n^2},$$

és így

$$\sin \xi(T_+^n x) \geq \frac{s_j(x)}{2n^2} \quad (n = 1, 2, \dots).$$

Bizonyítás. Mivel a gömbfelületen tekintett $T_+^n x$ pont $\frac{s_j(x)}{n^2}$ sugarú környezetében nincs szingularis pont, a $T_+^n x$ pontnak az északi és a déli saroktól való távolsága nagyobb vagy egyenlő, mint $\frac{s_j(x)}{n^2}$, ami állításunkat adja (19. ábra).



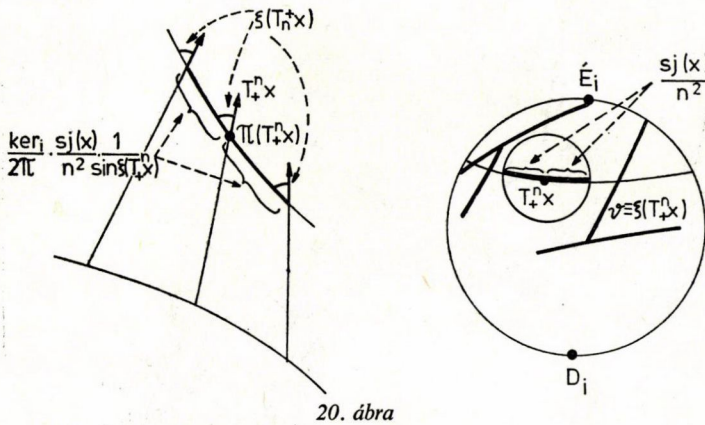
19. ábra

4.3 TÉTEL. Az a fésű, amelynek középpontja $\Pi(T_+^n x)$, sugara

$$\frac{\ker_i}{2\pi} \frac{sj(x)}{n^2} \frac{1}{\sin \xi(T_+^n x)},$$

nyoma pedig konstans, mégpedig $\xi(T_+^n x)$ -szel egyenlő, a T_- -transzformációnál nem szakad szét ($n-1, 2, \dots$). (Itt \ker_i annak az ütközőnek a kerületét jelenti, amelyiken a $T_+^n x$ elem van).

Bizonyítás. Mivel a gömbfelületen tekintett $T_+^n x$ pont $\frac{sj(x)}{n^2}$ sugarú környezetében nincs szinguláris pont, a $\mathfrak{S} \equiv \xi(T_+^n x)$ görbének a $T_+^n x$ pont mindkét oldalán legalább $\frac{sj(x)}{n^2}$ hosszúságú darabja nem tartalmazza T_+^{-1} -nek szakadási pontját, ami fésűk nyelvére lefordítva állításunkat igazolja (20. ábra).



20. ábra

5. Összesimuló sugárnyaláb konstrukciója

Legyen

$$B := \sqrt{A} = \exp\left(\frac{\omega}{2} \tau_{\min}\right).$$

Mivel véges sok ütköző van, $\min_i \{\ker_i\}$ pozitív konstans. Ezért $\exists D_1 > 0$ konstans, mely csak a biliárdasztaltól függ, hogy

$$D \cdot \frac{1}{B^{n-1}} \leq \frac{1}{2\pi} \min_i \{\ker_i\} \cdot \frac{1}{n^2} \quad (n = 1, 2, \dots; \quad 0 < D \leq D_1).$$

Vegyünk egy $x \in M$ elemet, melyre $sj(x) > 0$. $0 < sj(x) \leq 1$ miatt $sj^2(x) \leq sj(x)$. Ezért a 4.3. tételből adódik a következő tétel.

5.1. TÉTEL. $0 < D \leq D_1$ esetén az az $a_{n,D}$ fésű, melynek középpontja $\pi(T_+^n x)$, sugara

$$a_{n,D} := D \cdot \frac{sj^2(x)}{B^{n-1}} \cdot \frac{1}{\sin \xi(T_+^n x)},$$

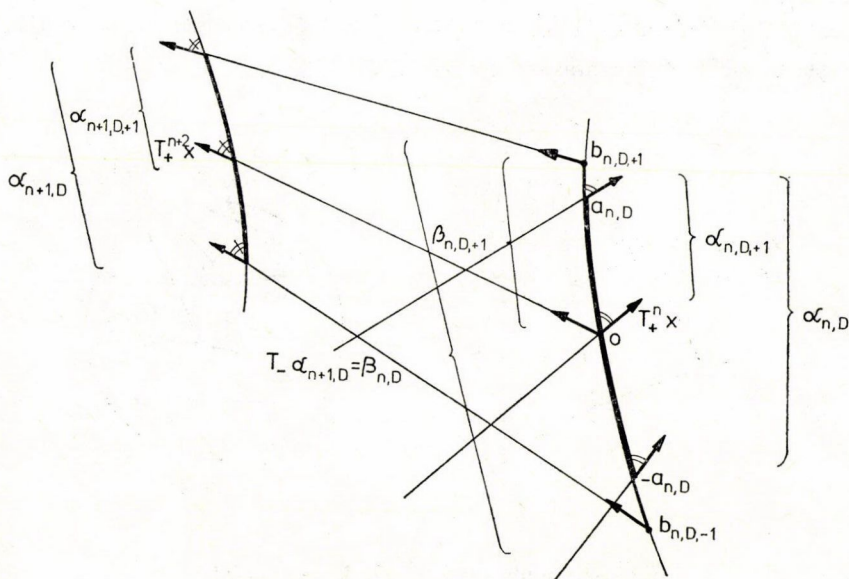
nyoma pedig kostans, mégpedig $\xi(T_+^n x)$ -szel egyenlő, a T_- -transzformációnál nem szakad szét ($n=1, 2, \dots$).

Vegyük észre, hogy az $\alpha_{n,D}$ fésű szélessége

$$W(\alpha_{n,D}) = 2D \cdot \frac{\text{sj}^2(x)}{B^{n-1}} \quad (n=1, 2, \dots; \quad 0 < D \leq D_1).$$

A következő tétel kapcsolatot teremt az $\alpha_{n,D}$ fésűk között.

5.2. TÉTEL. $\exists D_0 > 0$ konstans, mely csak a biliárdasztaltól függ, hogy $0 < D \leq D_0$ esetén $D(T_- \alpha_{n+1,D}) \supset D(\alpha_{n,D})$ ($n=1, 2, \dots$) (21. ábra).



21. ábra

Bizonyítás. $\alpha_{n,D}$ megszorítását a $[-a_{n,D}, 0]$, illetve a $[0, a_{n,D}]$ szakaszokra $\alpha_{n,D,-1}$, illetve $\alpha_{n,D,+1}$ jelöli. Legyen továbbá a $\beta_{n,D,i} := T_- \alpha_{n+1,D,i}$ fésű hossza $b_{n,D,i}$ ($i = \pm 1$). Igazolandó: $b_{n,D,i} \geq a_{n,D,i}$ ($i = \pm 1$). Az igazolást az $i = +1$ esetre végezzük el. Felhasználva a 3.8. tétel analógiáját kapjuk, hogy

$$\begin{aligned} \int_0^{b_{n,D,+1}} \sin \beta_{n,D,+1}(r) dr &= W(\beta_{n,D,+1}) \cong A \cdot W(\alpha_{n+1,D,+1}) = \\ &= \sqrt{A} \cdot W(\alpha_{n,D,+1}) = \sqrt{A} \cdot a_{n,D} \cdot \sin \xi(T_+^n x). \end{aligned}$$

Ezért $b_{n,D,+1} \geq a_{n,D}$ igazolásához elég belátni, hogy $r \in [0, a_{n,D}]$ -re $\sin \beta_{n,D}(r) \leq \sqrt{A} \cdot \sin \xi(T_+^n x)$. Mivel $\beta_{n,D}(0) = \xi(T_+^n x)$, ehhez elég belátni, hogy

$$a_{n,D} \cdot \max_{r \in [0, a_{n,D}]} \left| \frac{d(\sin \beta_{n,D}(r))}{dr} \right| \leq (\sqrt{A} - 1) \cdot \sin \xi(T_+^n x).$$

Az alábbi a), b) és c) pontokból kitűnik, hogy ez minden pozitív, egész n -re teljesül, ha D kisebb egy elég kicsire választott $D_0 > 0$ konstansnál.

a) A 4.2. tétel második egyenlőtlenségét felhasználva kapjuk, hogy

$$a_{n,D} = D \cdot \frac{\text{sj}^2(x)}{B^{n-1}} \cdot \frac{1}{\sin \xi(T_+^n x)} \leq 2D \cdot \text{sj}(x) \cdot \frac{n^2}{B^{n-1}}.$$

b)
$$\left| \frac{d(\sin \beta_{n,D}(r))}{dr} \right| = |\cos \beta_{n,D}(r)| \cdot \left| \frac{d\beta_{n,D}(r)}{dr} \right| \leq \Omega_1,$$

hiszen a (3.1) egyenlőtlenség analógja szerint $\beta_{n,D} = T_{-\alpha_{n+1,D}}$ fésű nyomának deriváltját abszolút értékben felülről korlátozza a jobb oldali konstans.

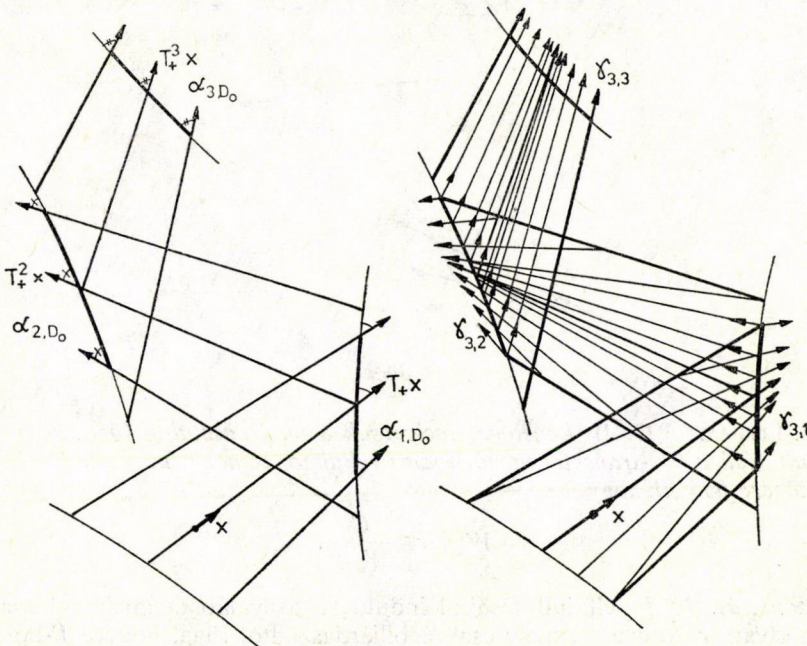
c)
$$(\sqrt{A}-1) \cdot \sin \xi(T_+^n x) \leq (\sqrt{A}-1) \cdot \frac{\text{sj}(x)}{2n^2}.$$

Az 5.2. tételben definiált D_0 értékkel minden olyan $x \in M$ elemhez, melyre $\text{sj}(x) > 0$, egy $\alpha_n := \alpha_{n,D_0}$ ($n=1, 2, \dots$) fésűsorozatot kapunk. Ennek felhasználásával a $\gamma_{n,i}$ ($i=n, \dots, 1$; $n=1, 2, \dots$) fésűket rögzített n mellett i szerinti indukcióval definiáljuk:

$$\gamma_{n,n} := \alpha_n$$

$$\gamma_{n,i} := T_{-\gamma_{n,i+1}} \text{ megszorítása } D(\alpha_i)\text{-re } (i = n-1, \dots, 1).$$

Az 5.2. tétel és a 3.15. következmény szerint $D(\alpha_i) = D(\alpha_{i,D_0}) \subset D(T_{-\alpha_{i+1,D_0}}) \subset D(T_{-\gamma_{n,i+1}})$, így az indukciós eljárás egyik lépésnél sem szakad meg (22. ábra).



22. ábra

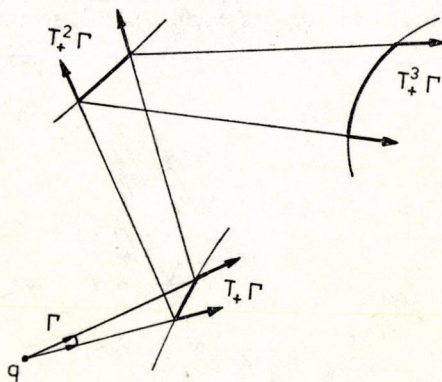
5.3. TÉTEL.

1. $\exists \gamma_i := \lim_{n \rightarrow \infty} \gamma_{n,i}$.
2. $A \gamma_i$ fésű nyoma folytonosan differenciálható, és $\gamma'_i = \lim_{n \rightarrow \infty} \gamma'_{n,i}$.
3. $A \gamma_i$ fésű a $T_{-\gamma_{i+1}}$ fésű megszorítása.
4. $A \gamma_i$ fésű középpontja $\pi(T_+^i x)$, sugara

$$D_0 \frac{sj^2(x)}{B^{i-1}} \sin \xi(T_+^i x) \quad (i = 1; 2, \dots).$$

A tétel bizonyításához egy lemmára van szükség. Először ezt fogjuk kimondani és bebizonyítani.

Tekintsük a biliárdasztalnak egy q pontját, és ebben a pontban az irányoknak egy szögtartományát. A szögtartományhoz tartozó irányok halmazát *elfajult fésűnek* nevezzük a q pontban. Tehát egy Γ elfajult fésű a q pontban részhalmaza $\pi^{-1}(q)$ -nak. A szögtartomány nagyságát a Γ elfajult fésű *nyílásszögének* nevezzük, és $W(\Gamma)$ -val jelöljük. Kézenfekvő, hogy egy elfajult fésű esetén mit jelent $T_+ \Gamma$, és hogy $T_+ \Gamma$ nyíló fésű. Nem kell részletezni a T_+ -transzformációnál szét nem szakadó elfajult fésű definícióját sem. A $T_+^n \Gamma$ fésűt n szerinti indukcióval definiáljuk $n > 1$ -re: ha $\Gamma, T_+ \Gamma, \dots, T_+^{n-1} \Gamma$ szét nem szakadó fésűk, akkor $T_+^n \Gamma := T_+(T_+^{n-1} \Gamma)$. Ilyenkor azt mondjuk, hogy a Γ elfajult fésű a T_+ -transzformáció n -szeri alkalmazásánál *nem szakad szét* (23. ábra).



23. ábra

5.4. LEMMA. $\exists C > 0$ konstans, mely csak a biliárdasztaltól függ, hogy ha egy Γ elfajult fésű a T_+ -transzformáció n -szeri alkalmazásánál *nem szakad szét*, akkor *nyílásszögére fennáll*, hogy

$$W(\Gamma) \leq \frac{C}{A^{n-1}}.$$

Bizonyítás. A Γ elfajult fésűből induló sugárnyaláb definíciója kézenfekvő. Létezik olyan r konstans, mely csak a biliárdasztaltól függ, hogy a Γ -ból induló sugárnyalábnak a q ponttól a sugarak mentén r távolságra levő ortogonális

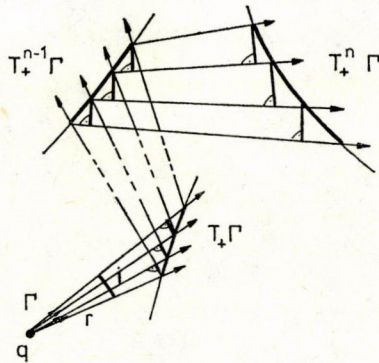
trajektóriájának i ívhosszára teljesül, hogy $W(\Gamma) \leq \frac{2i}{r}$. A 3.6. következmény és a 3.8. tétel miatt igazak az alábbi egyenlőtlenségek (24. ábra).

$$i \leq w(T_+ \Gamma),$$

$$w(T_+ \Gamma) \leq \frac{1}{A} \cdot w(T_+^2 \Gamma),$$

$$\vdots$$

$$w(T_+^{n-1} \Gamma) \leq \frac{1}{A} \cdot w(T_+^n \Gamma).$$



24. ábra

Nyilván fennáll, hogy $W(T_+^n \Gamma) \leq \ker_{\max}$, ahol $\ker_{\max} = \max_i \{\ker_i\}$ az ütközők területének maximuma. Mindebből kiadódik, hogy

$$W(\Gamma) \leq \frac{2 \cdot \ker_{\max}}{r} \cdot \frac{1}{A^{n-1}}.$$

Tehát a $C := \frac{2 \cdot \ker_{\max}}{r}$ konstans megfelel.

Az 5.3. tétel bizonyítása

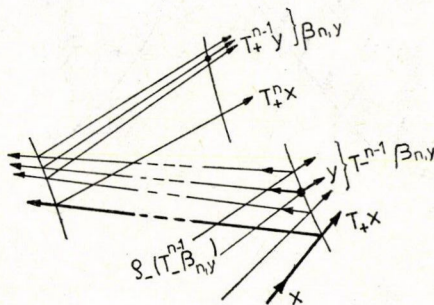
1. *Bizonyítása.* Az állítást elég $i=1$ -re igazolni. Legyen $\Delta := D(\alpha_1) = D(\alpha_1, d_0)$. Ha $q \in \Delta$, akkor $m, n \geq N$ esetén a $\gamma_{m,1}(q)$ és $\gamma_{n,1}(q)$ irányok által meghatározott szögtartományhoz tartozó elfajult fésű a T_+ -transzformáció N -szeri alkalmazásánál nem szakad szét, és ezért a $\gamma_{m,1}(q), \gamma_{n,1}(q)$ vektorok által bezárt szög kisebb vagy egyenlő, mint $\frac{c}{A^{N-1}}$. Mivel $N \rightarrow \infty$ esetén $\frac{c}{A^{N-1}} \rightarrow 0$, ebből következik, hogy a $\gamma_{n,1}$ ($n=1, 2, \dots$) sorozat egyenletesen konvergens Δ -n.

2. *Bizonyítása.* Ezt az állítást is elég $i=1$ -re igazolni. Ezért elég belátni, hogy a $\gamma'_{n,1}$ ($n=1, 2, \dots$) sorozat egyenletesen konvergens Δ -n, mert mint tudjuk, ebből következik, hogy γ_1 folytonosan differenciálható, és $\gamma'_1 = \lim_{n \rightarrow \infty} \gamma'_{n,1}$.

$q \in \Delta$ esetén legyen $\Gamma_n(q) \subset \pi^{-1}(q)$ az a maximális elfajult fésű q -ban, mely tartalmazza a $\gamma_{n,1}(q)$ elemet, és nem szakad szét a T_+ -transzformáció n -szeri alkalmazásánál. Mint tudjuk, $W(\Gamma_n(q)) \cong \frac{c}{A^{n-1}}$ ($n=1, 2, \dots$). Legyen továbbá $\Gamma_n := \bigcup_{q \in \Delta} \Gamma_n(q) \subset$

$\subset \partial M$ ($n=1, 2, \dots$), $\Gamma := \bigcap_{n=1}^{\infty} \Gamma_n \subset \partial M$. Nyilván a $\Gamma \cap \pi^{-1}(q)$ halmaz egyetlen eleme a $\gamma_1(q)$ elem ($q \in \Delta$).

$y \in \Gamma_n$ esetén legyen $\beta_{n,y}$ az a konstans nyomú fésű, mely a $\pi(T_+^{n-1}y)$ pontnak a biliárdasztal szélén vett környezetében értelmezett, $\beta_{n,y}(\pi(T_+^{n-1}y)) = T_+^{n-1}y$ és a T_- -transzformáció $(n-1)$ -szeri alkalmazása során nem szakad szét (25. ábra).



25. ábra

A $T_+^{n-1}\beta_{n,y}$ fésűbe érkező $\mathcal{L}(T_+^{n-1}\beta_{n,y})$ sugárnyaláb szóródása a $\pi(y)$ pontban legyen $\kappa_n(y)$ ($n=1, 2, \dots$). A Γ_n halmaz kompaktsága és a κ_n függvény folytonossága miatt κ_n egyenletesen folytonos ($n=1, 2, \dots$).

A 3.9. tételből következik, hogy $m \equiv n \equiv N$, $y \in \Gamma_m$ esetén

$$|\kappa_m(y) - \kappa_n(y)| \leq \frac{\Omega_0}{A^{2N-2}}.$$

Ezért $y \in \Gamma$ esetén $m \equiv n \equiv N$ -re teljesül, hogy

$$|\kappa_m(y) - \kappa_n(y)| \leq \frac{\Omega_0}{A^{2N-2}}.$$

Ez pedig azt jelenti, hogy a κ_n megszorítása Γ -n ($n=1, 2, \dots$) olyan függvénysorozat ad, mely egyenletesen tart egy $\kappa_\infty: \Gamma \rightarrow \mathbb{R}^1$ függvényhez. A $\kappa_n: \Gamma_n \rightarrow \mathbb{R}^1$ ($n=1, 2, \dots$) függvények egyenletes folytonosságából a kapott egyenlőtlenség felhasználásával könnyű látni, hogy a κ_n ($n=1, 2, \dots$) függvények egyenlő mértékben egyenletesen folytonosak.

Most e két tény felhasználásával igazoljuk, hogy a $\gamma'_{n,1}$ ($n=1, 2, \dots$) sorozat egyenletesen konvergens Δ -n. Egy csukódó fésűbe érkező sugárnyaláb szóródása és a fésű nyoma közötti kapcsolat alapján

$$\gamma'_{n,1}(q) = -k(q) - \kappa_n(\gamma_{n,1}(q)) \cdot \sin \gamma_{n,1}(q) \quad (q \in \Delta).$$

Ezért elég azt igazolni, hogy a $\kappa_n(\gamma_{n,1}(q)) \cdot \sin \gamma_{n,1}(q)$ ($n=1, 2, \dots$) sorozat $q \in \Delta$ -ra egyenletesen konvergál, a $\kappa_\infty(\gamma_1(q)) \cdot \sin \gamma_1(q)$ határértékhez.

Ez pedig kiolvasható az alábbiakból.

$$\begin{aligned} & |\kappa_\infty(\gamma_1(q)) \cdot \sin \gamma_1(q) - \kappa_n(\gamma_{n,1}(q)) \cdot \sin \gamma_{n,1}(q)| \leq \\ & \leq |\kappa_\infty(\gamma_1(q)) \cdot \sin \gamma_1(q) - \kappa_n(\gamma_1(q)) \cdot \sin \gamma_1(q)| + \\ & + |\kappa_n(\gamma_1(q)) \cdot \sin \gamma_1(q) - \kappa_n(\gamma_{n,1}(q)) \cdot \sin \gamma_{n,1}(q)|. \end{aligned}$$

A jobb oldalon álló első tag κ_n -nek κ_∞ -hez való egyenletes konvergenciája miatt elég nagy n -re kisebb ε -nál. A második tag pedig $\gamma_{n,1}$ -nek γ_1 -hez való egyenletes konvergenciája és a κ_n ($n=1, 2, \dots$) függvények egyenlő mértékben való egyenletes folytonossága miatt elég nagy n -re ε -nál kisebb, ahol az „elég nagy” kifejezés csak ε -tól függő, q -tól pedig nem függő küszöbindexre utal.

3. *Bizonyítás.* Ha egy $r \in D(\gamma_i)$ pontban $\gamma_i(r)$ nem volna egyenlő $T_{-\gamma_{i+1}}(r)$ -rel, akkor e két vektor egy szögtartományt határozna meg, amelyhez olyan elfajult fésű tartozna, ami a T_+ -transzformáció akárhányszori alkalmazásánál sem szakadna szét ($i=1, 2, \dots$).

4. *Bizonyítás.* Mivel $D(\gamma_i) = D(\alpha_{i,D_0})$, a γ_i fésű középpontja és sugara megegyezik az α_{i,D_0} fésű középpontjával és sugarával. Az 5.3. tétel bizonyításából az is kiolvasható, hogy

$$\gamma'_1(q) = -k(q) - \kappa_\infty(\gamma_1(q)) \cdot \sin \gamma_1(q) \quad (q \in \Delta).$$

Ezért a γ_1 fésűbe érkező sugárnyaláb szóródása egy $q \in \Delta$ helyen $\kappa_\infty(\gamma_1(q))$. Mivel $\kappa_\infty(\gamma_1(q)) = \lim_{n \rightarrow \infty} \kappa_n(\gamma_1(q))$, és $\kappa_n(\gamma_1(q)) \leq -k_{\min}$, a 3.1. tétel analógiájából kapjuk, hogy a γ_1 fésűbe érkező $\mathcal{S}_-(\gamma_1)$ sugárnyaláb szóródása a $\mathcal{P}_-(\gamma_1)$ halmazon mindenhol kisebb vagy egyenlő, mint $-\omega$. Természetesen a γ_i fésűbe érkező sugárnyalábról is állíthatjuk ugyanezt ($i=1, 2, \dots$). Így adódik a következő

5.5. TÉTEL. Az $\mathcal{S}_-(\gamma_i)$ sugárnyaláb szóródása a $\mathcal{P}_-(\gamma_i)$ halmazon mindenhol kisebb vagy egyenlő, mint $-\omega$ ($i=1, 2, \dots$).

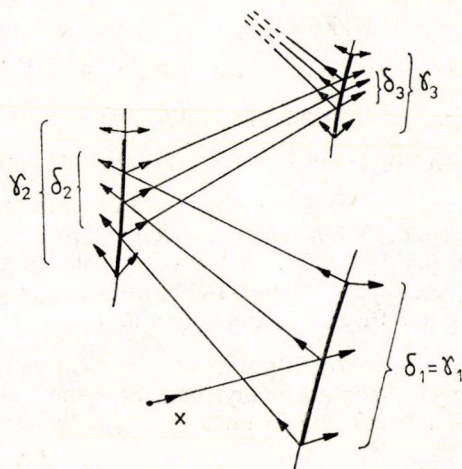
Az 5.3. tétel szerint minden olyan $x \in M$ elemhez, melyre $\text{sj}(x) > 0$, egy γ_i ($i=1, 2, \dots$) fésűsorozatot kapunk. Ennek felhasználásával definiáljuk most a δ_i fésűket:

$$\delta_1 := \gamma_1$$

δ_{i+1} legyen az a fésű, melyre $\delta_i = T_{-\delta_{i+1}}$ ($i=1, 2, \dots$). Az 5.3. tétel 3. állítása miatt ez a definíció korrekt, és az is látszik, hogy a δ_i fésű a γ_i fésű megszorítása ($i=1, 2, \dots$) (26. ábra). A 3.8. tétel analógiája szerint $W(\delta_i) \cong A \cdot W(\delta_{i+1})$ ($i=1, 2, \dots$), s így

$$(5.1) \quad W(\delta_i) \cong \frac{W(\delta_1)}{A^{i-1}} \quad (i=1, 2, \dots).$$

Annak a konstrukciónak az eredményeként, melyet egy $x \in M$ ($\text{sj}(x) > 0$) elcmből kiindulva ennek a szakasznak az elején kezdtünk el, most eljutottunk a $\delta_1, \delta_2, \dots$ csukódó fésűkhöz. Tekintsük a $\delta_1, \delta_2, \dots$ fésűkbe érkező $\mathcal{S}_-(\delta_1), \mathcal{S}_-(\delta_2), \dots$ sugárnyalábokat. Mivel $\delta_i = T_{-\delta_{i+1}}$ ($i=1, 2, \dots$), ezek között a sugárnyalábok között az alábbi kapcsolat áll fenn: $\mathcal{S}_-(\delta_{i+1})$ azokból a sugarakból áll, melyeket akkor kapunk, ha egy $\mathcal{S}_-(\delta_i)$ -beli sugáron haladó golyót a biliárdasztal szélén ütköz-



26. ábra

tetünk, és az ütközés utáni pályákat tekintjük a következő ütközésig ($i = 1, 2, \dots$). Az $\mathcal{S}_-(\delta_1), \mathcal{S}_-(\delta_2), \dots$ sugárnyalábokat együtt az x elemhez tartozó, ∞ -ben össze-simuló vagy szűkülő sugárnyalábnak nevezzük, és $\mathcal{S}^{(sz)}(x)$ -szel vagy röviden $\mathcal{S}^{(sz)}$ -szel jelöljük.

Az $\mathcal{S}_-(\delta_1), \mathcal{S}_-(\delta_2), \dots$ sugárnyalábok közötti, előbb elmondott kapcsolat alapján az $\mathcal{S}^{(sz)}$ szűkülő sugárnyaláb végtelen hosszú „fél-pályákból” tevődik össze. Ezeket a „fél-pályákat” az $\mathcal{S}^{(sz)}$ szűkülő sugárnyaláb *sugarainak* nevezzük. Az alábbi lemma elemi eszközökkel igazolható.

5.6. LEMMA. Legyen $a_i < 0 < b_i$ ($i = 1, 2, \dots$). Tegyük fel, hogy $f_i: [a_i, b_i] \rightarrow [0, \infty)$ folytonosan differenciálható függvény, melyre

- 1) $f_i(0) \equiv \frac{c_1}{i^2}$,
- 2) $|f'_i(r)| \equiv c_2 \quad (a_i \equiv r \equiv b_i)$,
- 3) $\int_{a_i}^{b_i} f_i(r) dr \equiv \frac{c_3}{d^{i-1}} \quad (i = 1, 2, \dots)$,

ahol $c_1, c_2, c_3 > 0, d > 1$. Ekkor $\exists i_0 = i_0(c_1, c_2, c_3, d)$, hogy

$$f_i(r) \equiv \frac{c_1}{2i^2} \quad (a'_i \equiv r \equiv b_i; i \geq i_0).$$

A δ_i fésű $D(\delta_i)$ értelmezési tartományát az ívhossz szerint azonosítsuk egy $[a_i, b_i]$ intervallummal úgy, hogy a $\pi(T_+^i x) \in D(\delta_i)$ pont a 0-nak feleljen meg. A kimondott lemmát alkalmazzuk az alábbi szereposztásban:

$$f_i(r) = \sin \delta_i(r) \quad (r \in D(\delta_i) \equiv [a_i, b_i], \quad i = 1, 2, \dots), \quad c_1 = \frac{1}{2} \text{sj}(x),$$

$$c_2 = \Omega_1, \quad c_3 = W(\delta_1), \quad d = A.$$

A lemma feltételei teljesülnek, mert

1. a 4.2. tétel második egyenlőtlensége szerint $f_i(0) = \sin \underline{\delta}_i(0) = \sin \xi(T_+^n x) \cong \frac{\text{sj}(x)}{2i^2} = \frac{c_1}{i^2}$,
2. a (3.1) egyenlőtlenség analógja szerint a $\delta_i = T_- \delta_{i+1}$ fésű nyomának deriváltját felülről korlátolja az Ω_1 konstans, és így $|f_i'(r)| = |\cos \underline{\delta}_i(r)| \cdot |\delta_i'(r)| \leq c_2$,
3. az (5.1) képlet szerint $\int_{a_i}^{b_i} f_i(r) dr = W(\delta_i) \leq \frac{W(\delta_1)}{A^{i-1}} = \frac{c_3}{d^{i-1}}$.

A lemma alapján igaz tehát a következő

5.7. TÉTEL. $\exists i_0 = i_0(x)$, hogy $i \geq i_0$ esetén $r \in D(\delta_i)$ -re $\sin \underline{\delta}_i(r) \geq \frac{\text{sj}(x)}{4i^2}$.

A tétel jelentősége abban rejlik, hogy eddig csak $\sin \underline{\delta}_i(0)$ -ra volt alsó becslésünk (4.2. tétel), most pedig már $\sin \underline{\delta}_i(r)$ -re is tudunk alsó becslést ($r \in D(\delta_i)$).

Az $\mathcal{S}_-(\delta_i)$ és az $\mathcal{S}_-(\delta_{i+1})$ sugárnyaláboknak a fentebb leírt kapcsolatából következik, hogy az $\mathcal{S}_-(\delta_i)$ sugárnyalábnak és az $\mathcal{S}_-(\delta_{i+1})$ sugárnyalábnak egy $r \in D(\delta_i)$ helyen vett szóródásainak különbsége $\frac{-2 \cdot k(r)}{\sin \underline{\delta}_i(r)}$ ($i = 1, 2, \dots$). A 3.3. tétel analógja miatt az $\mathcal{S}_-(\delta_{i+1})$ sugárnyaláb szóródása egy $r \in D(\delta_i)$ helyen nagyobb vagy egyenlő, mint $-\Omega_0$. Ezért az 5.7. tétel alapján $i \geq i_0$ -ra az $\mathcal{S}_-(\delta_i)$ sugárnyalábnak egy $r \in D(\delta_i)$ helyen a szóródása nagyobb vagy egyenlő, mint

$$-\Omega_0 - \frac{8 \cdot k_{\max}}{\text{sj}(x)} \cdot i^2 \geq -c(x) \cdot i^2,$$

ahol $c(x)$ x -től függő, de i -től nem függő konstans jelöl.

A 3.4. tétel analógja alapján ebből azt a következtetést vonhatjuk le, hogy $i \geq i_0$ -ra az $\mathcal{S}_-(\delta_i)$ sugárnyaláb szóródása a $\mathcal{P}_-(\delta_i)$ halmazon mindenhol nagyobb vagy egyenlő, mint

$$\min(-c(x) \cdot i^2; -\sqrt{|K_{\min}|}) \geq -d(x) \cdot i^2,$$

ahol $d(x)$ is x -től függő, de i -től nem függő konstans jelöl. Tehát kiadódik a következő

5.8. TÉTEL. $\exists i_0 = i_0(x)$, és $\exists d = d(x)$, hogy $i \geq i_0$ -ra az $\mathcal{S}_-(\delta_i)$ sugárnyaláb szóródása a $\mathcal{P}_-(\delta_i)$ halmazon abszolútértékben mindenhol kisebb vagy egyenlő, mint $d \cdot i^2$.

A tétel számunkra azért lesz hasznos, mert garantálja, hogy az $\mathcal{S}_-(\delta_i)$ sugárnyaláb szóródásának abszolútértéke i növekedtével csak négyzetes rendben növekedhet.

Az $\mathcal{S}^{(sz)}$ szűkülő sugárnyaláb sugarainak ortogonális trajektóriái segítségével a sugarakon vett ívhosszparamétereket összeegyeztethetjük úgy, hogy minden egyes ortogonális trajektória minden sugarat ugyanolyan paraméterű pontban messen, s így módon az ortogonális trajektória is egy paraméterértéknek feleljen meg. A t paraméterű ortogonális trajektóriát jelöljük g_t -vel (27. ábra).

Az 5.5. tétel miatt az $\mathcal{S}^{(sz)}$ szűkülő sugárnyaláb szóródása mindenhol kisebb vagy egyenlő, mint $-\omega$. Ezért a 3.7. következmény analógja miatt a g_t ortogo-

nális trajektóriák ívhosszaira fennáll, hogy

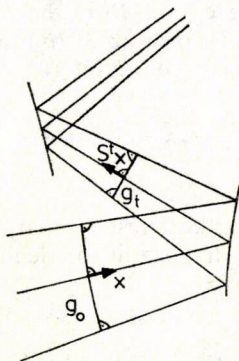
$$g_t \text{ ívhossza} \leq g_0 \text{ ívhossza} \cdot \exp(-\omega \cdot t) \quad (t \geq 0).$$

Jelöljük most \tilde{g}_t -vel azt a görbét a fázistérben, melyet úgy kapunk, hogy a g_t minden pontjára egy „ $\mathcal{S}^{(sz)}$ -irányú”, g_t -re merőleges egységvektort ültetünk. A (2.1) képlet szerint

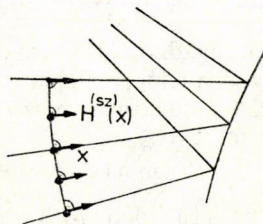
$$\tilde{g}_t \text{ ívhossza} = \int_{g_t} \sqrt{1 + (\kappa(t, q))^2} \, \varrho(dg)$$

ahol $\kappa(t, q)$ a g_t görbének a $q \in g_t$ pontban vett görbületét, tehát az $\mathcal{S}^{(sz)}$ sugárnyaláb szóródását jelenti. Mivel g_t ívhossza t függvényében exponenciális rendben csökken, $\max \{|\kappa(t, q)| : q \in \mathcal{P}_-(\gamma_i), q \in g_t\}$ pedig i függvényeként csak négyzetes rendben nő, kiadódik, hogy \tilde{g}_t ívhossza 0-hoz tart, ha $t \rightarrow \infty$. Ez a tény mutatja, hogy egy szűkülő sugárnyaláb tényleg „szűkül”, illetve „ ∞ -ben összesimul”.

Vegyük észre, hogy ha $\pi(x)$ pont a biliárdasztal belsejében van, és a $\pi(x)$ ponton átmenő ortogonális trajektória a t_0 paraméterértékhez tartozik, akkor a $H^{(sz)}(x) := g_{t_0}$ görbe rendelkezik mindazon tulajdonságokkal, melyeket az x elemhez tartozó szűkülő horociklustól a 2. szakaszban megköveteltünk (28. ábra). Ezzel megkaptuk ennek a szakasznak a fő eredményét:



27. ábra



28. ábra

5.9. TÉTEL. Ha az x elem a fázistér belsejében van, és $s_j(x) > 0$, akkor az x elemhez tartozik szűkülő horociklus.

Mivel a fázistér széle nyilván null-mértékű, a 4.1. és az 5.9. tételekből folyik, hogy a fázistér majdnem minden eleméhez tartozik szűkülő horociklus. Természetesen ennek az állításnak az analógja is igaz, ami végül is kiadja a 2. szakaszban kimondott egzisztencia-tétel állítását: a fázistér majdnem minden eleméhez tartozik szűkülő és táguló horociklus.

6. Szinaj-biliárd potenciáltérben

Tegyük fel, hogy a Szinaj-féle biliárdasztalon adott egy kétszer folytonosan differenciálható potenciálfüggvény, mely nem azonosan konstans. Szemléletesen ez azt jelenti, hogy a biliárdasztal nem „vízszintes”, hanem a potenciálfüggvénynek meg-

felelően „hepe-hupás”. Képzeljük el, hogy a potenciál hatása alatt egy pontszerű, egységnyi tömegű golyó mozog a biliárdasztalon. A golyó a potenciál „hepe-hupáinak” megfelelően görbe vonalú pályán, változó nagyságú sebességgel halad a biliárdasztalon, az ütközőkről pedig a rugalmas ütközés szabályai szerint visszaverődik.

A golyó állapotát a helye és a sebessége határozza meg. Ezért a biliárdasztal érintővektorait tekintjük állapotoknak. M_h -val jelöljük azoknak az állapotoknak a halmazát, melyekben az energia (potenciális energia + kinetikus energia) h -val egyenlő ($h \in \mathbb{R}^1$).

A golyó mozgása egy $\{Z'\}_{t \in \mathbb{R}^1}$ dinamikát értelmez az állapotok halmazán. Az energiamegmaradás törvénye miatt az M_h halmazok invariánsak erre a dinamikára.

Mint ismeretes (lásd például [4]-ben) az állapotok halmazán tekintett *Liouville-mérték* invariáns a $\{Z'\}_{t \in \mathbb{R}^1}$ dinamikára. Ezért a *Liouville-mérték* μ_h feltételes eloszlásai az M_h halmazokon ugyancsak invariánsak $\{Z_t\}_{t \in \mathbb{R}^1}$ -re. Ily módon minden egyes h energiaszinthez egy $(M_h, \{Z'\}_{t \in \mathbb{R}^1}, \mu_h)$ dinamikus rendszert kapunk, amit a *potenciál hatása alatt h energiaszinten vett Szinaj-biliárdnak* nevezünk.

Jelöljük U -val a potenciálfüggvényt. Rögzítsünk egy h energiaszintet. Az M_h halmaz a biliárdasztal bizonyos érintővektoraiból áll. Egy olyan érintővektor, mely a q ponthoz tartozó érintősíkban fekszik ($q \in Q$), M_h definíciója miatt, akkor és csak akkor eleme M_h -nak, ha hossza $\sqrt{2 \cdot (h - U(q))}$. Ezért ha minden M_h -beli érintővektort megfeleltetünk a vele egy érintősíkban levő és vele egy irányba mutató egységvektornak, akkor az M_h -t kölcsönösen és egyértelműen képezzük le a biliárdasztalt érintő egységvektorok M halmazára. Ezért az M halmazt is tekintjük az U potenciál hatása alatt h energiaszinten vett *Szinaj-biliárd fázisterének*.

Mint ismeretes (lásd [4]-ben a 60. oldalon), az U potenciál függvényhez és a h energiaszinthez a biliárdasztalon be lehet vezetni egy olyan Riemann-metrikát, hogy a megadott h energiaszinten mozgó golyók pályái a biliárdasztalon geodetikus vonalak legyenek ebben az új metrikában: ha $q \in Q$, és ds tetszőleges irányú, q -n átmenő „elemi görbedarab”, akkor az új *Riemann-metrikában* ds ívhossza

egyenlő az euklideszi *Riemann-metrikában* vett ívhossz $\sqrt{1 - \frac{U(q)}{h}}$ -szorosával.

Látszik, hogy az új *Riemann-metrika* szerint a biliárdasztal *Gauss-görbülete* nem azonosan nulla, vagyis az új *Riemann-metrika* szerint a biliárdasztal görbült. Tekintsük most ezen a görbült biliárdasztalon az 1. szakaszban definiált biliárdot! Ebben a biliárdban a golyó az új *Riemann-metrika* szerint egységnyi sebességgel halad geodetikus vonalak mentén. Ennek a biliárdnak a fázistere is nyilván azonosítható az M halmazzal.

Most vegyük figyelembe, hogy az új *Riemann-metrikát* éppen úgy vezettük be, hogy a potenciáltérben vett *Szinaj-biliárdban* a golyó pályája geodetikus vonalakból álljon. Igaz, a golyó nem egységnyi, hanem változó nagyságú sebességgel halad geodetikus vonalak mentén! Mégis nyilvánvaló, hogy a fázisterre vonatkozó, fentebbi megállapodásunk értelmében a két biliárd fázistere és trajektóriái — mint halmazok — egybeesnek, és a két biliárd ergodikus szempontból ekvivalens egymással.

IRODALOM

- [1] VETIER, A., „A Szinai-biliárd ergodicitása” a *Káosz* c. kötetben, szerk.: Szépfalusi P. és Tél T. (Akadémiai Kiadó, 1982) 479—502.
- [2] Синай, Я. Т., «Динамические системы с упругими отражениями», Успехи Математических Наук XXV/2. (1970) 141—192.
- [3] Арнольд, В. И., Математические методы классической механики, (Наука, Москва, 1979).
- [4] Корнфельд И. П., — Синай Я. Г., — Фомин С. В., Эргодическая теория (Наука, Москва. 1980.)

(Beérkezett: 1982. október 11.)

VETIER ANDRÁS
BME VILLAMOSMÉRNÖKKARI MATEMATIKA TANSZÉK
1111 BUDAPEST, EGRY JÓZSEF U. 20—22.

SINAI-BILLIARD IN POTENTIAL FIELD

A. VETIER

The ergodicity of the so called Sinai-billiard was proved in 1970 [2]. The proof has the following four main parts:

- a) construction of “stable and unstable fibers”,
- b) proof of “absolute-continuity”,
- c) application of the so called “Hopf-method”,
- d) proof of the “Main Lemma”.

In this paper, I examine the more general case when a potential field is given on the billiard-table. One can imagine that the billiard-table is not flat but there are “hills” on it. The aim of this paper is to construct stable and unstable fibers for this “hilly billiard”.

A POISSON-ELOSZLÁS JELENTŐSÉGE ÖSSZETETT KÉMIAI REAKCIÓK SZTOCHASZTIKUS MODELLJÉBEN

TÓTH JÁNOS, ÉRDI PÉTER és TÖRÖK LÁSZLÓ TURUL

Budapest

„Tekintettel arra a folyamatos vitára,
amely a kémiai reakciókban végbemenő
fluktuációkról zajlik, hasznosnak tűnik akár
csak egy világosan megfogalmazott ered-
mény is.”

VAN KAMPEN [19, 333. oldal]

Ismertetjük az összetett kémiai reakciók globális sztochasztikus modelljét a *reakciók Volpert-féle megadása* alapján. Felírjuk a modellre vonatkozó evolúciós egyenleteket.

Elégséges feltételt adunk arra, hogy egy reakció stacionárius eloszlása egycsúcsú legyen. Egyszerű születési-halálozási reakciók és polinomiális egyszerű populációs *Markov-folyamatok* esetén szükséges és elégséges feltételt adunk arra, hogy egy reakció stacionárius eloszlása *Poisson-eloszlás* legyen. Kiderül, hogy ez — az irodalom egy részében kifejezett várákozással ellentétben — igen ritkán teljesül.

Vázoljuk a termodinamikai fluktuációelméletnek és a stacionárius eloszlások alakjának a kapcsolatát.

Megfogalmazunk néhány, az alkalmazások szempontjából érdekesnek tűnő, nyílt matematikai problémát.

Külön függelékben ismertettük WHITTLE egy eredményét, amely bizonyos folyamatok stacionárius eloszlására vonatkozik.

Ez a dolgozat a [9, 10, 41, 45] angol nyelvű dolgozatok matematikai eredményeit foglalja össze.

1. Bevezetés

Az e helyen megjelent korábbi dolgozatainkban az összetett kémiai reakciók globális (tehát térbeli inhomogenitásokat figyelmen kívül hagyó), folytonos idejű, folytonos állapotterű, *determinisztikus* modelljével foglalkoztunk, lásd például [43]. Ez a determinisztikus modell — bár történetileg is, gyakorlatilag is fontos — nem elegendően pontos mindenfajta reakciókinetikai jelenség követésére. *Kis rendszerek* tanulmányozásánál, „*egzotikus*” *jelenségek* (instabilitás, multistacionaritás, periodicitás) vizsgálatánál, és *paraméterbecslésnél* esetenként hasznosabb lehet az összetett kémiai reakció szokásos, folytonos idejű, globális *sztochasztikus* modellje, amely a kémiai rendszer állapotát az atomi-molekuláris szemlélettel összhangban diszkrétnek tekinti, és figyelembe veszi a véletlen ingadozásokat, továbbá inkább kapcsolatba hozható a nemegyensúlyi termodinamika elméletével [6, 8. oldal; 7; 8].

Ezt a modellt elsőként LEONTOVICS fogalmazta meg 1935-ben. A negyvenes évek munkáiból elsősorban — a később fizikusból biokémikussá lett — DELBRÜCK, továbbá SIEGERT egy-egy dolgozata említendő meg.

LEONTOVICS, SIEGERT (és az alább idézendő GANS) nem kifejezetten kémiai komponensekről, hanem különböző energiaszintű állapotokról beszélnek, így mun-

kájuk (bár formálisan ugyanolyan modellre vonatkozik, mint amilyenről itt szó lesz) közvetlenül inkább a statisztikus mechanika tárgyköréhez kapcsolódik. DELBRÜCK egy biokémiai példán, a tripszinogénnek tripszinné való átalakulásán mutatja meg, hogy a folyamat kezdeti szakaszát szükséges és lehetséges a megfelelő sztochasztikus modellel tárgyalni. Ezen a területen ő volt az első, aki a diszkrét állapotterű modellt folytonos állapotterűvel közelítette.

Igen jelentős volt az elmélet és az alkalmazások szempontjából is a majdnem két tucatnyi magyar szerző hozzájárulása a területhez. RÉNYINEK az ionizáció és rekombináció problémáját tárgyaló, sokszor idézett dolgozata volt az első, amely *másodrendű* reakció sztochasztikus modelljének teljes tárgyalását adta. Ebből a példából kiderül, hogy a sztochasztikus modell várható értékére felírható differenciálegyenlet *nem* egyezik meg az ugyanazon reakcióra szokásos determinisztikus modellként felírt differenciálegyenlettel. PRÉKOPA és munkatársai pedig az elsők között mutattak példát a modell nemtriviális kémiai alkalmazására: ők hosszú láncmolekulák bomlásával foglalkoztak.

Ebben a dolgozatban azt vizsgáljuk meg, hogy milyen szerepe van a *Poisson-eloszlásnak* az összetett kémiai reakciók stacionárius viselkedésének a leírásában.

Miért foglalkozunk ezzel a kérdéssel?

A kémiai reakciókinetika a kémiai reakciók időbeli lefolyásával foglalkozik. Ez a jelenségkör speciális esete annak, amit a termodinamika tárgyal. Ennek ellenére a két tudományág egymástól majdnem teljesen függetlenül fejlődött. Egyre inkább terjed azonban az az igény, hogy a kémiai reakciókinetika eredményeit specializálással meg kell tudni kapni a termodinamikáiból, illetve, ahol szükséges, a termodinamikát módosítani kell úgy, hogy ez a levezetés lehetségessé váljék. Ennek szellemében teszünk most néhány megjegyzést, amely a fenti kérdésre válaszként szolgálhat, és erre fogunk visszatérni az eredmények ismertetése után azok diszkussziójánál.

A kémiai reakciókinetika csaknem egy évszázadon keresztül determinisztikus modellekkel dolgozott, amelyek nemcsak az ingadozások természetéről nem mondanak semmit, hanem még az ingadozások létéről sem vesznek tudomást, ugyanakkor már a korai termodinamikai fluktuációelmélet számára is nyilvánvaló volt az (egyensúlyi vagy nemegyensúlyi) ingadozások ténye. A klasszikus termodinamika — így a fluktuációelmélet is — többnyire *folytonos* makroszkopikus változókkal dolgozik, és ezek ingadozását a legtöbb esetben normális eloszlással írja le [12, 26]. A *Poisson-eloszlás*, mint a normális eloszlást közelítő *egycsúcsú* eloszlás került be a termodinamikai fluktuációelméletbe [26, 397. oldal].

A sztochasztikus reakciókinetika meglehetősen bőséges irodalmában elterjedt az a nézet, hogy az összetett kémiai reakciók stacionárius eloszlása általában *Poisson-eloszlás* [37]. Ez valóban igaz ki- és beáramlást is tartalmazó, elsőrendű elemi reakciókból álló rendszerekre, más néven (pontosabban) nyílt rekeszrendszerekre [13, 692. oldal]. Ha a nyílt rekeszrendszerek modelljét úgy tekintjük, mint tetszőleges összetett kémiai reakció „egyensúlyhoz közeli” leírását, *ebben a közelítésben az állítás tetszőleges folyamatra igaz*. Emellett szól továbbá az a tény is, hogy bizonyos feltételek mellett a *Poisson-folyamatnak* az entrópiája a maximális [39] s ez a tulajdonság összhangban van a termodinamika II. főtételel. Ennek a ténynek a fontosságát azonban gyengíti az a körülmény, hogy számos nevezetes eloszláshoz létezik olyan feltétel, amely mellett éppen az adott eloszlás maximalizálja az entrópiát

[17, 1. oldal], és a feltételek korántsem következnek egyértelműen a feladat jellegéből [27, R109. oldal].

A kérdéskör irodalmát tovább bonyolította az a tény, hogy foglalkozni kezdtek az összetett kémiai reakció lokális (azaz a komponensek térbeli diffúzióját is figyelembe vevő, és így a termodinamikához közelebb álló) sztochasztikus modelljével — ezekről egy figyelemre méltó dolgozat [1] — miközben a globális modell tulajdonságai még nem voltak teljesen felderítve (ahogyan még ma sincsenek). Összefoglalva az irodalmat, azt mondhatjuk: elképzelhetőnek tartják, hogy a *Poisson-eloszlás* a diffúzió eltorzítja, de akkor egy olyan modellt kell definiálni, amely a diffúziót is le tudja írni, és az fog *Poisson-eloszlásra* vezetni.

Ebben a dolgozatban, figyelembe véve az előző bekezdésekben mondottakat és a mottót, a globális modell tanulmányozására szorítkozunk, azaz azokat a strukturális (kémiai kifejezéssel élve: *sztochiometriai*) feltételeket vizsgáljuk, amelyek mellett egy összetett kémiai reakció *Poisson-eloszlásra* vezethet.

A dolgozat felépítése a következő:

A 2. fejezetben ismertetjük az összetett kémiai reakciók sztochasztikus modelljére vonatkozó evolúciós egyenleteket. A 3. fejezetben elégséges feltételt adunk a stacionárius eloszlás egycsúcsúságára, illetve arra, hogy az *Poisson-eloszlás* legyen. A 4. fejezetben rávilágítunk az eredmények termodinamikai hátterére és számos megoldatlan problémát vetünk fel. Külön függelékben ismertetjük WHITTLE [47] egy tételét, amely bizonyos folyamatok stacionárius eloszlásának létezésére és egyértelműségére vonatkozik.

Az *ugró Markov-folyamatokra* vonatkozó alapvető állítások például a [4, 10A, 18, 32A] könyvekben találhatók meg. Felhívjuk az Olvasó figyelmét a [18] könyvre, amely az alkalmazásokra irányuló könyvek közül a legmagasabb színvonalú és a legbősebb.

2. Az összetett kémiai reakciók (globális) sztochasztikus modellje

2.1. A sztochasztikus modell és a kombinatorikus sztochasztikus modell

Először [43] 2.6. és 2.9. definíciója alapján megmondjuk, hogy mit értünk egy (tömeghatás kinetikájú) összetett kémiai reakció (*Volpert-féle*) megadásán. Az egyes objektumok jelentését itt nem részletezzük, lásd erről például a [42, 43] dolgozatokat.

2.1. DEFINÍCIÓ. *Összetett kémiai reakciónak* nevezzük az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ objektumot, ahol

(i) \mathcal{S} és \mathcal{R} egy-egy M , illetve $R (M, R \in \mathbb{N})$ elemű, egymástól diszjunkt halmaz, elemeiket *kémiai komponenseknek*, illetve *elemi reakcióknak* hívjuk;

(ii) $(\alpha, \beta) \in (\mathbb{N}_0 \times \mathbb{N}_0)^{\mathcal{S} \times \mathcal{R}}$ egy függvény, értékei a *sztochiometriai együtthatók*, amelyek megmutatják, hogy az egyes komponensekből hány darab szerepel az elemi reakciókat szokásosan jelképező nyíl bal, illetve jobb oldalán, más szóval reaktánsként, illetve terméként);

(iii) $k: \mathcal{R} \rightarrow \mathbb{R}^+$ egy olyan függvény, amelyik mindegyik elemi reakcióhoz hozzárendeli az elemi reakció *sebességi állandóját*.

Vezessük be a következő jelölést: tekintsük az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ összetett kémiai reakciót, és legyen

$$(2.1) \quad \mathcal{R}_x := \{r \in \mathcal{R}; \beta(., r) - \alpha(., r) = x\} \quad (x \in \mathbb{Z}^{\mathcal{S}}).$$

Legyen továbbá

$$(2.2) \quad R' := |\{\mathcal{R}_x; 0 \neq x \in \mathbb{Z}^{\mathcal{S}}\}|.$$

Nyilvánvaló, hogy

- (i) $R' \leq R$,
- (ii) $\mathcal{R}_x \cap \mathcal{R}_{x'} = \emptyset$, ha $x, x' \in \mathbb{Z}^{\mathcal{S}}$ és $x \neq x'$;
- (iii) $\bigcup \{\mathcal{R}_x; 0 \neq x \in \mathbb{Z}^{\mathcal{S}}\} = \mathcal{R}$.

Feltehető, hogy $\mathcal{R}_0 = \emptyset$ (itt $0 \in \mathbb{Z}^{\mathcal{S}}$).

2.2. DEFINÍCIÓ. Az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ összetett kémiai reakció (indukált) sztochasztikus modellje egy olyan $\xi: \Omega \times \mathbb{R}_0^+ \rightarrow \mathbb{N}^{\mathcal{S}}$ (folytonos idejű, időben homogén, diszkrét állapotterű Markov-) folyamat, amelynek infinitezimális mátrixában

$$(2.3) \quad a_{jl} := \sum_{r \in \mathcal{R}_{l-j}} k(r) j^{\alpha(., r)} \quad (j, l \in \mathbb{N}_0^{\mathcal{S}}; j \neq l),$$

a többi elem pedig 0.

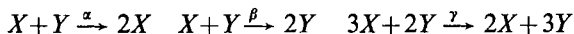
2.3. DEFINÍCIÓ. Az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ összetett kémiai reakció (indukált) Kurtz-féle vagy kombinatorikus sztochasztikus modellje egy olyan $\xi: \Omega \times \mathbb{R}_0^+ \rightarrow \mathbb{N}_0^{\mathcal{S}}$ (folytonos idejű, időben homogén, diszkrét állapotterű Markov-) folyamat, amelynek infinitezimális mátrixában

$$(2.4) \quad a_{jl}^K := \sum_{r \in \mathcal{R}_{l-j}} k(r) [j]_{\alpha(., r)} \quad (j, l \in \mathbb{N}_0^{\mathcal{S}}; j \neq l),$$

a többi elem pedig 0. (Itt $[j]_{\alpha(., r)} := \prod_{m=1}^M j_m(j_m - 1) \dots (j_m - \alpha(m, r) + 1)$ minden olyan $j \in \mathbb{N}_0^M$ esetére, amelyre $j_m \geq \alpha(m, r)$ minden $m \in \{1, 2, \dots, M\}$ mellett.

Mielőtt általános állításokat mondanánk ki, hasznos lehet részletesen megvizsgálni egy példát. A kombinatorikus modellre vonatkozó adatokat egy felső K indexszel különböztetjük meg, a q_1, q_2, q_1^K, q_2^K függvényekre később lesz szükségünk.

2.1. PÉLDA: Tekintsük az



összetett kémiai reakciót! Ennél

$$\mathcal{S} = \{X, Y\}; \mathcal{R} = \{1, 2, 3\}; k(1) = \alpha, k(2) = \beta, k(3) = \gamma,$$

$$\alpha(., 1) = (1, 1)^T \quad \alpha(., 2) = (1, 1)^T \quad \alpha(., 3) = (3, 2)^T$$

$$\beta(., 1) = (2, 0)^T \quad \beta(., 2) = (0, 2)^T \quad \beta(., 3) = (2, 3)^T$$

$$\mathcal{R}_{(1, -1)^T} = \{1\} \quad \mathcal{R}_{(-1, 1)^T} = \{2, 3\} \quad R' = 2$$

$$a_{j, j+(1, -1)^T} = \alpha j_1 j_2 =: q_1(j)$$

$$a_{j, j+(-1, 1)^T} = \beta j_1 j_2 + \gamma j_1^2 j_2^2 =: q_2(j)$$

$$a_{jl} = 0 \quad \text{egyébként.}$$

Továbbá:

$$\begin{aligned} a_{j,j+(1,-1)}^K &= \alpha j_1 j_2 =: q_1^K(j), \\ a_{j,j+(-1,1)}^K &= \beta j_1 j_2 + \gamma j_1(j_1-1)(j_1-2)j_2(j_2-1) = \gamma j_1^3 j_2^2 - \gamma j_1^3 j_2 - 3\gamma j_1^2 j_2^2 + \\ &+ 3\gamma j_1^2 j_2 + 2\gamma j_1 j_2^2 + (\beta - 2\gamma)j_1 j_2 =: q_2^K(j). \end{aligned}$$

2.1. TÉTEL. Egy összetett kémiai sztochasztikus és kombinatorikus sztochasztikus modellje között

$$a_{jl} = a_{jl}^K \quad (j, l \in \mathbb{N}_0^S; j \neq l)$$

pontosan akkor áll fenn, ha $\alpha(m, r) \leq 1$ minden m és r esetén.

Bizonyítás: Az állítás a 2.2. és 2.3. definíciónak közvetlen következménye.

2.2. *Megjegyzés.* (i) Az $\alpha(m, r) \leq 1$ feltétel [43] 4.6. definíciója szerint azt jelenti, hogy a reakció a *gyengén reálisak* közé tartozik. Az idézett fogalom általánosítása a *reális* összetett kémiai reakciónak, amelynél $\sum_{m=1}^M \alpha(m, r) \leq 2$, vagyis amelynél a reakciók legfeljebb másodrendűek. A reális reakciók viszont éppen a legfontosabbak a kémiai alkalmazások szempontjából, amit azzal szokás indokolni, hogy az ilyenek végbemenéséhez szükséges kettős ütközések még lehetségesek, míg a hármas és többszörös ütközések már nem.

(ii) A tétel érdekessége éppen abban áll, hogy a kétféle sztochasztikus modell azonosságának feltétlenül egy olyan tulajdonságot ad meg, amely tulajdonság idáig a determinisztikus modellel kapcsolatban látszott csak fontosnak (ti. annak eldöntésénél, hogy egy determinisztikus modell gradiens rendszer-e).

(iii) Itt is megfogalmazható és megoldható a [43]-ban tárgyalt inverz feladat analogonja, az eredmény is hasonló, lásd [42, 81–84. oldal].

2.2. Az összetett kémiai reakció sztochasztikus modelljére vonatkozó evolúciós egyenletek

2.2.1. Az átmenetvalószínűségekre vonatkozó Kolmogorov-egyenletek

A 2.2. definícióból behelyettesítéssel kapjuk az átmenetvalószínűségekre vonatkozó *Kolmogorov-egyenleteket*.

2.1. ÁLLÍTÁS. Az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ reakció sztochasztikus modelljének átmenetvalószínűségeire fennáll:

$$\begin{aligned} \dot{P}_{jl}(t) &= \sum_{k \in \mathbb{N}_0^S} \sum_{r \in \mathcal{R}_{k-j}} k(r) j^{\alpha(\cdot, r)} P_{kl}(t) - P_{jl}(t) \sum_{k \in \mathbb{N}_0^S} \sum_{r \in \mathcal{R}_{k-j}} k(r) j^{\alpha(\cdot, r)}, \\ (2.5) \quad \dot{P}_{jl}(t) &= \sum_{k \in \mathbb{N}_0^S} P_{jk}(t) \sum_{r \in \mathcal{R}_{1-k}} k(r) k^{\alpha(\cdot, r)} - P_{jl}(t) \sum_{k \in \mathbb{N}_0^S} \sum_{r \in \mathcal{R}_{k-1}} k(r) l^{\alpha(\cdot, r)}, \\ P_{jl}(0) &= \delta_{jl} \quad (j, l \in \mathbb{N}_0^S). \end{aligned}$$

2.2.2. Az abszolút valószínűségekre vonatkozó alapegyenlet, és a stacionárius eloszlásra vonatkozó egyenlet

Ismét behelyettesítéssel kapjuk az alapegyenletet, és a stacionárius eloszlásra vonatkozó egyenletet.

2.2. ÁLLÍTÁS. Az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ reakció sztochasztikus modelljének abszolút valószínűségeire teljesül, hogy:

$$(2.6) \quad \dot{P}_j(t) = \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} P_l(t) \sum_{r \in \mathcal{R}_{j-l}} k(r) l^{\alpha(\cdot, r)} - P_j(t) \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} \sum_{r \in \mathcal{R}_{l-j}} k(r) j^{\alpha(\cdot, r)} \quad (j \in \mathbb{N}_0^{\mathcal{S}}),$$

továbbá, ha a stacionárius eloszlást $\{\pi_j; j \in \mathbb{N}_0^{\mathcal{S}}\}$ jelöli, akkor

$$(2.7) \quad 0 = \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} \pi_l \sum_{r \in \mathcal{R}_{j-l}} k(r) l^{\alpha(\cdot, r)} - \pi_j \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} \sum_{r \in \mathcal{R}_{l-j}} k(r) j^{\alpha(\cdot, r)} \quad (j \in \mathbb{N}_0^{\mathcal{S}}).$$

2.2.3. A generátorfüggvényre és a momentumgeneráló függvényre vonatkozó egyenlet

2.4. DEFINÍCIÓ. Legyen a ξ folyamat állapottere \mathbb{N}_0^M . A folyamat generátorfüggvényének nevezzük a

$$\mathcal{G}(z, t) := \sum_{j \in \mathbb{N}_0^M} z^j P_j(t)$$

képlettel definiált, a

$$\mathcal{D} := \{(z, t) \in \mathbb{C}^M \times \mathbb{R}_0^+; |pr_m z| \leq 1, m \in \{1, 2, \dots, M\}\}$$

halmazon értelmezett függvényt, momentumgeneráló függvényének pedig a \mathcal{D} halmazon definiált,

$$\mathcal{M}(z, t) := \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} \exp(j^T z) P_j(t)$$

összefüggéssel megadott \mathcal{M} függvényt.

2.1. PÉLDA (folytatás). Legyenek a fenti \mathcal{D} halmazon definiált, differenciálható függvényeken a parciális deriválások operátorai $\partial_1, \partial_2, \dots, \partial_M$ (és ∂_{M+1}) és jelöljük \tilde{q}_1 -mal és \tilde{q}_2 -mal az alábbi leképezéseket:

$$\tilde{q}_1: (\mathcal{G}^K, z, t) \mapsto \alpha z_1 z_2 \partial_1 \partial_2 \mathcal{G}^K(z, t)$$

$$\tilde{q}_2: (\mathcal{G}^K, z, t) \mapsto \beta z_1 z_2 \partial_1 \partial_2 \mathcal{G}^K(z, t) + \gamma z_1^3 z_2^3 \partial_1^3 \partial_2^3 \mathcal{G}^K(z, t).$$

A példából látható a q_1 és \tilde{q}_1 , valamint a q_2 és \tilde{q}_2 közötti kapcsolat. A jelen keretek között ezt nem érdemes általánosan definiálni. Hasonlóan, legyen

$$\tilde{\tilde{q}}_1: (\mathcal{M}^K, z, t) \mapsto \alpha \partial_1 \partial_2 \mathcal{M}^K(z, t)$$

$$\tilde{\tilde{q}}_2: (\mathcal{M}^K, z, t) \mapsto \beta \partial_1 \partial_2 \mathcal{M}^K(z, t) + \gamma \partial_1^3 \partial_2^3 \mathcal{M}^K(z, t).$$

Itt a kapcsolat könnyen megadható:

$$\tilde{q}_1(\mathcal{M}^K, z, t) = [q_1((\partial_1, \partial_2)^T) \cdot \mathcal{M}^K](z, t),$$

$$\tilde{q}_2(\mathcal{M}^K, z, t) = [q_2((\partial_1, \partial_2)^T)] \cdot \mathcal{M}^K(z, t).$$

(Most a q_1 , illetve q_2 jelölést használtuk a megfelelő polinomok operátor-argumentumra való kiterjesztettjére is.)

2.2. TÉTEL. Legyen az $\langle \mathcal{S}, \mathcal{R}, (\alpha, \beta), k \rangle$ reakció kombinatorikus sztochasztikus modelljének generátorfüggvénye \mathcal{G}^K , momentumgeneráló függvénye pedig \mathcal{M}^K ; $q_{r'}$ ($r' \in \{1, 2, \dots, R'\}$) pedig legyen a sztochasztikus modell alapján — a példában leírt módon — definiálva. Akkor

$$(2.8) \quad \partial_{M+1} \mathcal{G}^K(z, t) = \sum_{r'=1}^{R'} (z^{x(r')} - 1) \tilde{q}_{r'}(\mathcal{G}^K, z, t)$$

$$(2.9) \quad \partial_{M+1} \mathcal{M}^K(z, t) = \sum_{r'=1}^{R'} (e^{x(r')^T z} - 1) \tilde{\tilde{q}}_{r'}(\mathcal{M}^K, z, t)$$

ahol $\tilde{q}_{r'}$ és $\tilde{\tilde{q}}_{r'}$ a 2. példában leírt módon származnak $q_{r'}$ -ből; $x(r') \in \mathbb{Z}^{\mathcal{S}}$ az r' -höz tartozó megváltozás, ∂_{M+1} pedig a hagyományosan $\partial/\partial t$ -vel jelölt operátor.

Bizonyítás. Csak (2.8)-at bizonyítjuk, (2.9) hasonló módon bizonyítható.

Szorozzuk meg (2.6) mindkét oldalát z^j -vel és összegezzünk j -re. Ha ezután azonos átalakításokat végzünk felhasználva az alapegyenletet, felcserélve az összegezés sorrendjét, megváltoztatva az indexezést, becsempészve a generátorfüggvény deriváltjait, végül felhasználva q_r definícióját, a következő egyenlőségsorozatot kapjuk:

$$\begin{aligned} \partial_{M+1} \mathcal{G}^K(z, t) &= \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} z^j \dot{P}_j(t) = \\ &= \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} z^j \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} P_l(t) \sum_{r \in \mathcal{A}_{j-l}} k(r)[l]_{\alpha(\cdot, r)} - \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} z^j P_j(t) \sum_{l \in \mathbb{N}_0^{\mathcal{S}}} \sum_{r \in \mathcal{A}_{l-j}} k(r)[j]_{\alpha(\cdot, r)} = \\ &= \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} \sum_{r \in \mathcal{A}} k(r) z^{\beta(\cdot, r)} z^{j-\beta(\cdot, r)} \cdot P_{j-\beta(\cdot, r)+\alpha(\cdot, r)}(t) [j-\beta(\cdot, r)+\alpha(\cdot, r)]_{\alpha(\cdot, r)} \\ &\quad - \sum_{r \in \mathcal{A}} z^{\alpha(\cdot, r)} \sum_{j \in \mathbb{N}_0^{\mathcal{S}}} k(r)[j]_{\alpha(\cdot, r)} z^{j-\alpha(\cdot, r)} = \sum_{r \in \mathcal{A}} k(r) z^{\beta(\cdot, r)} \left(\frac{\partial}{\partial z} \right)^{\alpha(\cdot, r)} \mathcal{G}^K(z, t) \\ &\quad - \sum_{r \in \mathcal{A}} k(r) z^{\alpha(\cdot, r)} \left(\frac{\partial}{\partial z} \right)^{\alpha(\cdot, r)} \mathcal{G}^K(z, t) = \sum_{r'=1}^{R'} \sum_{r \in \mathcal{A}_{x(r')}} k(r) z^{\alpha(\cdot, r)+x(r')} \left(\frac{\partial}{\partial z} \right)^{\alpha(\cdot, r)} \mathcal{G}^K(z, t) \\ &\quad - \sum_{r'=1}^{R'} \sum_{r \in \mathcal{A}_{x(r')}} k(r) z^{\alpha(\cdot, r)} \left(\frac{\partial}{\partial z} \right)^{\alpha(\cdot, r)} \mathcal{G}^K(z, t) = \sum_{r'=1}^{R'} z^{x(r')} \tilde{q}_{r'}(\mathcal{G}^K, z, t) \\ &\quad - \sum_{r'=1}^{R'} \tilde{q}_{r'}(\mathcal{G}^K, z, t) = \sum_{r'=1}^{R'} (z^{x(r')} - 1) \tilde{q}_{r'}(\mathcal{G}^K, z, t). \end{aligned}$$

2.3. Megjegyzés. (i) A tétel állításai skaláris esetben speciális esetei BAILEY [2, 7. fejezet] ugró Markov-folyamatokra vonatkozó állításainak.

(ii) Megjegyezzük, hogy a tétel a *kombinatorikus* sztochasztikus modell generátorfüggvényeire vonatkozó összefüggéseket ad meg, de ezekben a („közönséges”) sztochasztikus modell jellemzői szerepelnek.

(iii) Amint a bizonyításból kitűnik, fennáll a

$$(2.10) \quad \partial_{M+1} \mathcal{G}^K(z, t) = \sum_{r \in \mathcal{R}} k(r) (z^{\beta(\cdot, r)} - z^{\alpha(\cdot, r)}) \left(\frac{\partial}{\partial z} \right)^{\alpha(\cdot, r)} \mathcal{G}^K(z, t)$$

egyenlet is.

(iv) Egy további evolúciós egyenlet található [42] 4.1. tételében.

3. Egyensúlyi eloszlások

3.1. Az egycsúcsúság egy elégséges feltétele

Ebben a szakaszban olyan reakciók egyensúlyi eloszlásának egycsúcsúságával foglalkozunk, amelyekben egy belső komponens van, és amelyek sztochasztikus modellje egy születési-halálozási folyamat.

Megjegyezzük, hogy az egycsúcsúság nemcsak fizikai-kémiai szempontból érdekes, hanem számítástechnikai szempontból is hasznos.

3.1. DEFINÍCIÓ. [32, 144. oldal]: A $\{\pi_j; j \in \mathbb{N}_0\}$ eloszlás egycsúcsú, ha a

$$\pi_1 - \pi_0, \quad \pi_2 - \pi_1, \quad \pi_3 - \pi_2, \quad \dots$$

sorozatban pontosan egy előjelváltás van, azaz

$$\exists! a \in \mathbb{N}_0 [\pi_{a+1} - \pi_a < 0 \leq \pi_a - \pi_{a-1}]$$

$$\forall b \in \mathbb{N}_0 (b < a \Rightarrow (\pi_b - \pi_{b-1} \geq 0)) (\pi_{-1} := 0).$$

MEDGYESSY egy tételét [32, 146. oldal] általánosítva elégséges feltételt adunk arra, hogy egy diszkrét eloszlás mikor lesz egycsúcsú.

3.1. LEMMA. Tegyük fel, hogy

$$\gamma: \mathbb{N}_0 \rightarrow \mathbb{R}^+ \quad \text{és} \quad \delta: \mathbb{N}_0 \rightarrow \mathbb{R}^+$$

olyan függvények, amelyekre

$$(3.1) \quad \gamma(j) + \delta(j) \geq 1 \quad (j \in \mathbb{N}_0)$$

teljesül, és tegyük fel, hogy a $\{\pi_j; j \in \mathbb{N}_0\}$ eloszlás olyan, hogy

$$(3.2) \quad \pi_j \geq \gamma(j)\pi_{j+1} + \delta(j)\pi_{j-1}.$$

Ekkor $\{\pi_j; j \in \mathbb{N}_0\}$ egycsúcsú.

Bizonyítás. Megmutatjuk, hogy ha $a \in \mathbb{N}_0$ -ban az eloszlás megváltozása előjelet vált, akkor a -tól kezdve szigorúan monoton csökken. (A legkisebb ilyen $a \in \mathbb{N}_0$ az eloszlás csúcsa.) Tegyük fel tehát, hogy

$$\pi_{a+1} - \pi_a < 0 \leq \pi_a - \pi_{a-1} \quad (\pi_{-1} := 0).$$

Ekkor (3.2)-t felhasználva kapjuk, hogy

$$\begin{aligned}\pi_{a+2} - \pi_{a+1} &\leq \frac{\pi_{a+1} - \delta(a+1)\pi_a}{\gamma(a+1)} - \pi_{a+1} = \\ &= (\pi_{a+1} - \pi_a) \frac{\delta(a+1)}{\gamma(a+1)} + \frac{1 - \gamma(a+1) - \delta(a+1)}{\gamma(a+1)} \pi_{a+1} < 0.\end{aligned}$$

3.1. TÉTEL. Tegyük fel, hogy egy összetett kémiai reakcióban $M=1$, és hogy a reakció sztochasztikus modellje születési-halálózási folyamat

$$\psi(j) := \psi_0, \quad \text{illetve} \quad \mu(j) := \sum_{i=1}^l \mu_i j^i$$

$$j \in \mathbb{N}_0; \quad \psi_0 \in \mathbb{R}^+; \quad \mu_i \in \mathbb{R}_0^+, \quad i \in \{1, \dots, l\}, \quad l \in \mathbb{N}$$

születési, illetve halálózási aránnyal. Ekkor a reakció sztochasztikus modelljének létezik egyetlen stacionárius eloszlása, és az egysúcsú.

1. Bizonyítás. A stacionárius eloszlás létezése és egyértelműsége a *Karlin—McGregor-feltétel* [21] teljesüléséből következik. A reakció stacionárius eloszlására a (2.7) képlet alapján:

$$\begin{aligned}(3.3) \quad (\mu(j) + \psi(j))\pi_j &= \mu(j+1)\pi_{j+1} + \psi(j-1)\pi_{j-1} \\ (\pi_{-1} &:= 0, \quad j \in \mathbb{N}_0),\end{aligned}$$

ahonnan átrendezéssel, továbbá ψ speciális alakját és μ monotonitását felhasználva:

$$\begin{aligned}\pi_j &= \pi_{j+1} \frac{\mu(j+1)}{\mu(j) + \psi(j)} + \pi_{j-1} \frac{\psi(j-1)}{\mu(j) + \psi(j)} \cong \\ &\cong \pi_{j+1} \frac{\mu(j)}{\mu(j) + \psi(j)} + \pi_{j-1} \frac{\psi_0}{\mu(j) + \psi(j)}.\end{aligned}$$

Alkalmazva a 3.1. lemmát a

$$\gamma(j) := \frac{\mu(j)}{\mu(j) + \psi(j)}, \quad \delta(j) := \frac{\psi_0}{\mu(j) + \psi(j)} \quad (j \in \mathbb{N}_0)$$

szereposztással a bizonyítandó állítást kapjuk.

2. Bizonyítás. Születési-halálózási folyamatról lévén szó (3.3) mellett — amint az teljes indukcióval könnyen belátható —

$$(3.4) \quad \mu(j)\pi_j = \psi(j-1)\pi_{j-1} \quad (j \in \mathbb{N})$$

is teljesül. (3.4)-ből pedig ezt kapjuk:

$$(3.5) \quad \pi_j - \pi_{j-1} = \pi_j \frac{\psi(j-1) - \mu(j)}{\psi(j-1)} = \pi_j \frac{\psi_0 - \mu(j)}{\psi_0}$$

ahonnan a μ függvény szigorú monotonitásának felhasználásával látható, hogy a

$$N_0 \in j \mapsto \pi_j - \pi_{j-1} \quad (\pi_{-1} := 0)$$

sorozatban pontosan egy előjelváltás van.

3.1. *Megjegyzés.* A 2. bizonyításból kitűnik, hogy $R_\psi \subset \mathbb{R}^+$ esetén a

$$(3.6) \quad N_0 \ni j \mapsto h(j) := \psi(j-1) - \mu(j) \in \mathbb{R}$$

összefüggéssel definiált h függvény előjelváltásainak száma határozza meg $\{\pi_j; j \in N_0\}$ csúcseinak számát, ezért a 3.1. tétel általánosabban is megfogalmazható lett volna.

3.2. *Elégéses feltételek arra, hogy a stacionárius eloszlás Poisson-eloszlás legyen*

3.2.1. *Egyszerű születési-halálozási reakciók*

3.2. TÉTEL. [9]. Ha egy egyszerű születési-halálozási folyamatnak létezik stacionárius eloszlása, akkor az pontosan akkor *Poisson-eloszlás*, ha a folyamat lineáris, azaz ha az (F.2) és (F.3) képletben szereplő φ_m ($m \in \{1, 2, \dots, M\}$) homogén lineáris függvény.

Bizonyítás. Az F.1. tételből következik, hogy ha létezik stacionárius eloszlás, akkor az

$$(3.7) \quad \pi_j = \text{konst.} \cdot \frac{\alpha^j}{\prod_{m=1}^M \varphi_m(1) \dots \varphi_m(j_m)} \quad (j \in N_0^M)$$

alakú, ahol $\alpha \in \mathbb{R}^m$ az (F.6) egyenletrendszer megoldása. Teljes indukcióval látható, hogy a (3.7) eloszlás pontosan akkor

$$(3.8) \quad \pi_j = \exp \left(- \sum_{m=1}^M \lambda_m \right) \frac{\lambda^j}{j!}$$

$$(j \in N_0^M; \lambda \in (\mathbb{R}^+)^M; \lambda_m := \text{pr}_m \lambda; m \in \{1, 2, \dots, M\})$$

alakú, ha φ_m ($m \in \{1, 2, \dots, M\}$) homogén lineáris függvény.

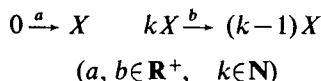
3.1. *Következmény.* Ha egy összetett kémiai reakció sztochasztikus modellje egyszerű születési-halálozási folyamat, és a folyamatnak létezik stacionárius eloszlása, akkor az pontosan akkor *Poisson-eloszlás*, ha $\sum_{m=1}^M \alpha(m, r) \leq 1$ minden r esetén.

3.2. *Megjegyzés.* (i) A tételben tehát a „szorzatformula” [22] egy speciális esetét használtuk fel.

(ii) Amint (3.8)-ból kitűnik, vektoriális folyamatok esetén akkor nevezzük a stacionárius eloszlást *Poisson-eloszlásnak*, ha koordinátái független *Poisson-eloszlású* valószínűségi változók.

(iii) A következményben szereplő feltételt úgy is szokás mondani, hogy ekkor a reakció nyílt *rekeszrendszer* [43]. Maga a következményben megfogalmazott állítás összhangban áll GANS [13, 692. oldal] eredményével.

3.1. PÉLDA. A



reakció, bár sztochasztikus modellje *egyszerű* születési-halálozási folyamat, a tétel szerint általában nem vezet *Poisson-eloszlásra*, kivéve a $k=1$ esetet. A $k=1$ -re vonatkozó speciális állításnak egy másik általánosításaként megvizsgáltuk, hogy azoknak a reakcióknak a sztochasztikus modellje, amelyeké egydimenziós születési-halálozási folyamat, mikor vezet *Poisson-eloszlásra* [45]. Később ezt az eredményt sikerült vektoriális folyamatokra általánosítani a részletes kiegyensúlyozottság (egydimenziós esetben automatikusan teljesülő, l. a 3.1. tétel 2. bizonyítását) feltételének feltevése mellett. Ezt az eredményt ismertetjük most [41] alapján.

3.2.2. Polinomiális egyszerű populációs Markov-folyamatok

3.2. DEFINÍCIÓ. Legyen

$$M \in \mathbf{N}; \quad a_m, b_m, c_{mn} \in \mathbf{N}_0;$$

$$\psi_{m,k} \in \mathbf{R}_0^+ (k \in \{0, 1, \dots, a_m\}), \quad \mu_{m,k} \in \mathbf{R}_0^+ (k \in \{0, 1, \dots, b_m\}),$$

$$\varphi_{mn,k} \in \mathbf{R}_0^+ (k \in \{0, 1, \dots, c_{mn}\}) (m, n \in \{1, 2, \dots, M\}; m \neq n)$$

és tegyük fel, hogy ha a_m vagy b_m pozitív, akkor ψ_{m,a_m} is (illetve μ_{m,b_m} is). Defináljuk a $\psi_m; \mu_m$ és φ_{mn} polinomokat $m, n \in \{1, 2, \dots, M\}; m \neq n$ esetén a következőképpen:

$$(3.9) \quad \psi_m(j_m) := \sum_{k=0}^{a_m} \psi_{m,k} j_m^k$$

$$(3.10) \quad \mu_m(j_m) := \sum_{k=0}^{b_m} \mu_{m,k} j_m^k$$

$$(3.11) \quad \varphi_{mn}(j_m) := \sum_{k=0}^{c_{mn}} \varphi_{mn,k} j_m^k$$

($j_m \in \mathbf{N}_0$; j_m helyett index nélküli változót is írhattunk volna, de ilyen alakban fogjuk használni a képleteket.) Tegyük fel, hogy a $\xi: \Omega \times \mathbf{R}_0^+ \rightarrow \mathbf{N}_0^M$ folyamat infinitezimális mátrixának nullától különböző elemeit a most definiált polinomok adják meg a következőképpen:

$$(3.12) \quad a_{j,j+e_m} = \psi_m(j_m),$$

$$(3.13) \quad a_{j,j-e_m} = \mu_m(j_m),$$

$$(3.14) \quad a_{j,j-e_m+e_n} = \varphi_{mn}(j_m)$$

j_m a $j \in \mathbb{N}_0^M$ vektor m -edik koordinátája: $j_m := pr_m j$. Ekkor a ξ folyamatot *polinomiális egyszerű populációs Markov-folyamatnak* hívjuk.

3.3. DEFINÍCIÓ. A $\xi: \Omega \times \mathbb{R}_0^+ \rightarrow \mathbb{N}_0^M$ folyamat *sztochasztikusan átmenetenként részletesen kiegyensúlyozott*, ha a $\{\pi_j; j \in \mathbb{N}_0^M\}$ stacionárius eloszlásra teljesül, hogy

$$(3.15) \quad a_{jl} \pi_j = a_{lj} \pi_l \quad (j, l \in \mathbb{N}_0^M).$$

3.3. TÉTEL. A ξ sztochasztikusan átmenetenként részletesen kiegyensúlyozott polinomiális egyszerű populációs Markov-folyamat stacionárius eloszlása pontosan akkor

$$(3.16) \quad \pi_j = \frac{\lambda^j}{j!} e^{-\sum_{m=1}^M \lambda_m} \quad (j \in \mathbb{N}_0^M)$$

alakú (ahol $\lambda \in (\mathbb{R}^+)^M$; $\lambda_m := pr_m \lambda$, ha teljesülnek a következő relációk:

$$(3.17) \quad a_m + 1 = b_m, \quad c_{mn} = 1, \quad \varphi_{mn,0} = 0;$$

$$(3.18) \quad \lambda_m = \psi_{m,a_m} / \mu_{m,a_m+1};$$

$$(3.19) \quad \psi_{m,r} = (\psi_{m,a_m} / \mu_{m,a_m}) \sum_{k=r+1}^{a_m+1} \binom{k-1}{r} \mu_{mk} \\ (r \in \{0, 1, \dots, a_m-1\});$$

$$(3.20) \quad \mu_{m,0} = 0;$$

$$(3.21) \quad \varphi_{nm,1} = \varphi_{mn,1} \frac{\psi_{m,a_m} \mu_{n,a_n+1}}{\psi_{n,a_n} \mu_{m,a_m+1}}$$

$$(m, n \in \{1, 2, \dots, M\}; \quad m < n).$$

Bizonyítás. Az elégségesség magától értendő lévén csak a szükségességet bizonyítjuk. A részletes kiegyensúlyozottság (amelyet *ugró Markov-folyamatok* esetében *reverzibilitásnak* is neveznek) esetünkben a következőt jelenti:

$$(3.22) \quad \psi_m(j_m) \pi_j = \mu_m(j_m+1) \pi_{j+e_m},$$

$$(3.23) \quad \varphi_{mn}(j_m) \pi_j = \varphi_{nm}(j_m+1) \pi_{j-e_m+e_n}$$

$$(j \in \mathbb{N}_0^M; \quad m, n \in \{1, 2, \dots, M\}).$$

(3.9)–(3.11)-et és (3.16)-ot behelyettesítve (3.22)–(3.23)-ba azonnal kapjuk (3.17)-et. A behelyettesítéskor a két oldalon kapott polinomok együtthatóinak összevetéséből

ez adódik:

$$(3.24) \quad \psi_{m,0} = \lambda_m (\mu_{m,0} + \dots + \mu_{m,a_m+1})$$

$$\psi_{m,0} + \psi_{m,1} = \lambda_m \left(\binom{1}{0} \mu_{m,1} + \binom{2}{1} \mu_{m,2} + \dots + \binom{a_m+1}{a_m} \mu_{m,a_m+1} \right)$$

...

$$(3.25) \quad \psi_{m,k-1} + \psi_{m,k} = \lambda_m \left(\binom{k}{0} \mu_{m,k} + \binom{k+1}{1} \mu_{m,k+1} + \dots + \binom{a_m+1}{a_m+1-k} \mu_{m,a_m+1} \right)$$

...

$$\psi_{m,a_m-1} + \psi_{m,a_m} = \lambda_m \left(\binom{a_m}{0} \mu_{m,a_m} + \binom{a_m+1}{1} \mu_{m,a_m+1} \right)$$

$$(3.26) \quad \psi_{m,a_m} = \lambda_m \mu_{m,a_m+1}$$

$$(3.27) \quad \varphi_{nm,1} \lambda_n = \varphi_{mn,1} \lambda_m \quad (m < n).$$

Azt mondhatjuk tehát, hogy van egy $\sum_{m=1}^M (a_m+2) + \binom{M}{2}$ számú egyenletből álló rendszerünk a következő

$$M + \sum_{m=1}^M (a_m+2) + \sum_{m=1}^M (a_m+1) + M(M-1) = 2 \sum_{m=1}^M a_m + M^2 + 3M$$

számú változóra:

$$\lambda_m; \mu_{m,a_m+1}, \dots, \mu_{m,0}; \psi_{m,a_m}, \dots, \psi_{m,0}; \varphi_{mn,1} \\ (m, n \in \{1, 2, \dots, M\}).$$

Tekintsük ismeretleneknek a következőket:

$$\lambda_m; \text{ és, ha } a_m > 0, \text{ akkor } \psi_{m,a_m-1}, \dots, \psi_{m,0}; \text{ és } \varphi_{nm,1} (m < n).$$

λ_m kifejezhető (3.26)-ból:

$$\lambda_m = \psi_{m,a_m} / \mu_{m,a_m+1}.$$

Nyilvánvaló, hogy a fennmaradó egyenletrendszernek létezik egyetlen pozitív megoldása a $\psi_{m,r} (r \in \{0, 1, \dots, a_m-1\})$ és $\varphi_{nm,1} (m < n; m, n \in \{1, 2, \dots, M\})$ változókban. Felhasználva, hogy (3.25) fennáll minden $m \in \{1, 2, \dots, M\}$ mellett, indukcióval kapjuk (3.19)-et. A λ_m -re és $\psi_{m,0}$ -ra kapott kifejezéseket (3.24)-be írva adódik (3.20). (3.21) pedig egyszerű átrendezettje (3.27)-nek.

3.2. PÉLDA. A

$$kA(m) \xrightarrow[\mu_{m,k+1}]{\psi_{m,k}} (k+1)A(m) \quad k \in \{0, 1, \dots, a_m\}$$

$$kA(m) \xrightarrow{\varphi_{mn,k}} A(n) + (k-1)A(m) \quad k \in \{0, 1, \dots, c_{mn}\} \\ (m, n \in \{1, 2, \dots, M\}).$$

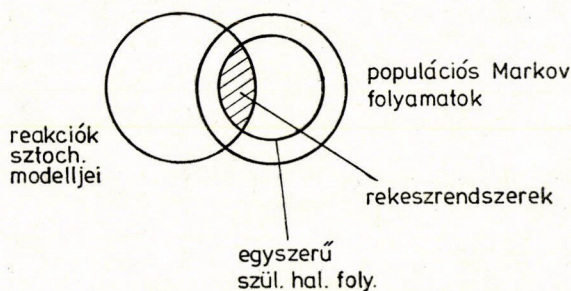
összetett kémiai reakció tehát pontosan akkor vezet Poisson-eloszlásra, ha a megfelelő sebességi állandók összefüggnek, és a közöttük fennálló összefüggést (3.18),

(3.19) és (3.21) adja meg. Abban a speciális esetben, amikor $M=1$, $c_{11}=0$, $a_1=0$, a

$$0 \xrightleftharpoons[\mu]{\psi} A$$

reakciót kapjuk, amelynél a megkötések nem jelentenek megszorítást a sebességi állandókra; ennek tehát mindig Poisson-eloszlás a stacionárius eloszlása.

3.3. *Megjegyzés.* Érdemes kiemelni, hogy a 3.2. tételben szereplő egyszerű születési-halálozási folyamat, és a populációs *Markov-folyamatok* (lásd pl. [23]) közé tartozó 3.3. tételbeli folyamat közötti lényeges különbség abban van, hogy az elsőnél a születési arányok az állapottól függetlenül állandók. Az egyes típusok közötti kapcsolatokat mutatja az 1. ábra.



1. ábra

4. Diszkusszió, nyílt problémák

4.1. Az egyensúly körüli ingadozásokról általában

Legyen $\mathbf{R} \ni t \mapsto \alpha(t)$ a tanulmányozandó folytonos idejű, időben homogén, \mathbf{R}^M állapotterű stacionárius *Markov-folyamat* ($M \in \mathbb{N}$ a termodinamikai változók száma). Rögzített $t \in \mathbf{R}$ esetén az $\alpha(t)$ valószínűségi vektorváltozó sűrűségfüggvénye — a folyamat abszolút sűrűségfüggvénye — legyen g . Ez a folyamatot gyakorlatilag kielégítően jellemzi; termodinamikai folyamatokra megadható az (alkalmasan definiált) entrópia függvényében.

A függvény alakja termodinamikai egyensúlyban:

$$(4.1) \quad g(x) = C \cdot \exp(-S(x)/k),$$

ahol a k a Boltzmann-állandó, C pedig egy normáló tényező. A (4.1) összefüggést *Einstein-képlet* néven szokás említeni. Az S függvényt az x^* érték körül — amelyről feltesszük, hogy termodinamikai egyensúlyi pont, vagyis, hogy S -nek stacionárius pontja — sorbafejtve, és a sorfejtésben a második tagnál megállva kapjuk, hogy a sűrűségfüggvényre közelítőleg

$$(4.2) \quad g(x) = \left(\frac{\det S''(x^*)}{2\pi k} \right)^{M/2} \exp\left[-\frac{1}{2k}(x-x^*)^T S''(x^*)(x-x^*)\right]$$

teljesül.

Matematikai szempontból a normális eloszlással, pontosabban a *Gauss-folyamattal* való közelítés mindig elég jó abban az értelemben, hogy minden első és második momentummal rendelkező sztochasztikus folyamathoz létezik olyan *Gauss-folyamat*, amelynek első és második momentuma megegyezik az eredetivel. Igaz továbbá, hogy az „átlagérték” és a „legvalószínűbb érték” (a várható érték és a módusz) megegyezik egymással (sőt még a mediánnal is).

A fluktuációt az entrópia felhasználása nélkül is ki szokás fejezni:

$$(4.3) \quad g(x) = \left(2\pi \sum_{m=1}^M x_m^* \right)^{-M/2} \exp\left(-\frac{1}{2}(x-x^*)^T (\text{dg } x^*)^{-1} (x-x^*)\right).$$

(Ez a [26] 396. oldalán szereplő képlet általánosítása *független* koordinátájú *vektor*-változókra; x^* koordinátáit x_1^*, \dots, x_M^* jelöli, $\text{dg } x^*$ pedig az x^* -ből képezett diagonális mátrix.)

(4.3)-at kis térfogatú rendszerekben, tehát viszonylag nagy fluktuációk esetén

$$(4.4) \quad P(N) = \frac{(x^*)^N e^{-N}}{N!} := \frac{\prod_{m=1}^M (x_m^*)^{N_m} \exp\left(-\sum_{m=1}^M N_m\right)}{\prod_{m=1}^M N_m!}$$

($N \in \mathbb{N}_0^M$)

alakú *Poisson-eloszlással* szokás közelíteni. Itt tulajdonképpen a valószínűségszámításban megszokott, normális eloszlással való közelítést fordítva alkalmazzuk.

4.2. Különböző szintű sztochasztikus reakciókinetikai modellek

KURAMOTO [24] bírálja NICOLIS és PRIGOGINE azon állítását, amely szerint a sztochasztikus kinetikai modell ellentétes a (*Boltzmann-egyenlet*en alapuló) makroszkopikus leírással. NICOLIS és munkatársai KURAMOTO rövid cikkére adott terjedelmes válaszukban [33] azt állítják, hogy KURAMOTO feltevései nem természetesek, és hogy a sztochasztikus kinetikai leírás bizonyos mikroszkopikus ingadozások leírására nem alkalmas. Valóban a sztochasztikus kinetikai modell átmenetet jelent az elemi ütközésekkel operáló „mikroszkopikus reakciókinetika” és a bruttó mennyiségekkel dolgozó fenomenologikus reakciókinetika között. Amíg a mikroszkopikus kinetika egyedi molekulákkal foglalkozik, addig a szokásosan megfogalmazott fenomenologikus reakciókinetika a koncentráció időfüggését tárgyalja. A koncentráció fogalmához a mikroszkopikus leírásból valamilyen várható érték-képzéssel juthatunk. Nyilvánvaló, hogy a szórás és a magasabb momentumok figyelembevétele a kinetikai folyamatoknak a fenomenologikus és determinisztikus kinetikánál részletesebb tanulmányozására ad lehetőséget. A sztochasztikus reakciókinetika — minthogy a leírás szintje átmenetet jelent a szigorúan mikroszkopikus és makroszkopikus között — VAN KAMPEN [20, 31. oldal] szerint: *mezozkopikus*.

A *Kuramoto-vita* kapcsán olyasféle álláspont alakult ki, mely szerint a kémiai fluktuációk lokálisan *Poisson-eloszlásra*, globálisan pedig ettől eltérő eloszlásra vezetnek.

A vita azért furcsa, mert részvevői *három különböző* modell tulajdonságait vetették össze. A mikroszkopikus, illetve a globális sztochasztikus modellen kívül a vizsgálatba bevonták a — nem is teljesen pontosan megfogalmazott — lokális sztochasztikus modellt is.

A globális és a lokális kinetikai modell nyilván csak igen részlegesen vethető össze. Determinisztikus modellek esetén senki sem kérné számon egy közönséges differenciálegyenlet-rendszertől (amelyben az ismeretlen függvények értelmezési tartományának elemeit *időpontokként* interpretáljuk), hogy nem írja le a *térbeli* változásokat. A szokásos sztochasztikus leírás viszont a determinisztikus modell parciális differenciálegyenlet-rendszerének nem egy véletlen mezőt, hanem egy sztochasztikus folyamatot feleltet meg. Másrésztől a klasszikus statisztikus mechanikában persze jól ismert, hogy a koncentráció-ingadozások térbelileg is korreláltak [26] és manapság — különösen a kritikus jelenségek modern, *Wilson-féle elméletével* [11, 48] kapcsolatban — igen fontos, hogy az egyes térrészek kis fluktuációi a térbeli kölcsönhatások eredményeképpen felerősödnek és makroszkopikusan drasztikus effektusokat okoznak.

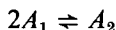
Álláspontunk az, hogy a mikroszkopikus kinetika figyelembevétele is és a térbeli korreláció is hathat abban az irányban, hogy az ingadozások „*Poisson-jellegét*” elrontja. Egyelőre érdemesnek tartottuk, hogy matematikai tételket mondjunk ki arra vonatkozólag, hogy az összetett kémiai reakció globális, folytonos idejű, diszkrét állapotterű sztochasztikus modelljének stacionárius eloszlása mikor *Poisson-eloszlás*. Most rátérünk annak ismertetésére, hogy hogyan kapcsolódnak eredményeink mások munkáihoz.

4.3. A Poisson-eloszlás előfordulása összetett kémiai reakcióknál

Cikkünk mottójának szerzője, VAN KAMPEN [19] abból a fizikai feltevésből indul ki, hogy a részecskeszám nagykanonikus eloszlása [26, 126. oldal] *Poisson-eloszlás*. Ebből — igen erős korlátozó feltevések mellett — megmaradási törvények figyelembe vételével kapja a stacionárius eloszlást. Erről az eloszlásról egyrészt kimutatja, hogy pontos megoldása az összetett kémiai reakció stacionárius eloszlására vonatkozó *alapegyenletnek* (*master equation*), másrészt, hogy még nagy rendszerek esetén is lényegesen különbözik a *Poisson-eloszlástól*.

HANUSSE [16, 86—88. oldal] felhívja a figyelmet arra, hogy ha az összetett kémiai reakcióban nemcsak monomolekuláris elemi reakciók vannak jelen, akkor előfordulhat, hogy a stacionárius eloszlás nem *Poisson-eloszlás*.

ZUREK és SCHIEVE [49] egy adott összetett kémiai reakció sztochasztikus modelljét szimulálva mutatja meg, hogy a stacionárius eloszlás nem *Poisson-eloszlás*. (Néhány kritikai észrevétel dolgozatunkhoz: olyan reakciót szimulálnak, amelynek stacionárius eloszlása kiszámítható, tehát a szimulálás eredménye csak az lehetne, hogy a programjuk nem hibás, viszont a szimuláció részletes leírásából kiderül, hogy nem azt a modellt szimulálják, amelyről az irodalomban szó van, hanem egy részletesebbet, fizikaibbat.) Mindezekhez képest elhanyagolható, hogy azt állítják, hogy a stacionárius eloszlás általában csak az elemi reakcióvektoroktól függ, maguktól az elemi reakcióktól nem; tehát például a



és a

$$3A_1 \rightleftharpoons A_1 + A_2$$

reakció stacionárius eloszlása azonos. A speciális állítás ugyan valóban igaz, de nagyon lényegesen kihasználja a vizsgált reakciók speciális szerkezetét.)

Végül megemlítjük, hogy a *Poisson-eloszlásnak* az irodalomban betöltött fontos szerepére mutat az is, hogy gyakran használják ezt az eloszlást vonatkoztatási alapnak. Nemlineáris rendszerek esetén a stacionárius eloszlás σ^2 szórásnégyzetét a stacionárius eloszlás várható értékének, m -nek a felhasználásával (skaláris folyamatról van szó!)

$$\sigma^2 = m \cdot (1 - \mu)$$

alakban írják fel, és μ -t a reakcióra jellemző paraméternek tekintik [16, 89. oldal; 30].

A *Poisson-eloszlás* vonatkoztatási alapként szerepel [6] 4.3. szakaszában, és a [14] dolgozatban is.

4.4. Nyílt problémák

Definiálni kellene az egycsúcsúság fogalmát a többdimenziós esetben, és a p -csúcsúság fogalmát ($p \in \mathbb{N}$) (legalább) egydimenziós esetben, majd ezekre MEDGYESSY eredményeinek általánosításával elégséges feltételeket kellene kimondani. Ezután ezeket az elégséges feltételeket alkalmazni kellene összetett kémiai reakciók stacionárius (vagy nem csak stacionárius) eloszlására. (Vö.: [10].)

Ezzel a problémakörrel szoros kapcsolatban áll az a kérdés, hogy mikor vezet egy reakció *diszkrét Pearson-eloszlásra* [36, 240. oldal], ugyanis ezek egycsúcsúsága könnyen vizsgálható.

A *Poisson-eloszlással* kapcsolatban felvethető, hogy melyek azok a reakciók, amelyek a sebességi állandókra tett külön megszorítások nélkül is ilyen egyensúlyi eloszlással bírnak, mint például a



reakció.

Kérdés az is, hogy a (megszorítással, vagy anélkül) *Poisson-típusú* egyensúlyi eloszlással rendelkező reakciók determinisztikus modelljéről tudunk-e mondani valamit. A megfelelő determinisztikus modell alkalmas jellemzése ugyanis valószínűleg közelebb visz bennünket a legfontosabb kérdés megválaszolásához, azéhoz tudniillik, hogy van-e valamilyen *fizikai-kémiai oka* annak, ha egy reakció stacionárius eloszlása *Poisson-eloszlás*?

Végül ismét egy olyan kvalitatív inverz feladatot említünk, amelynek *fizikai háttere* jelentős: Melyek azok a reakciók, amelyek stacionárius eloszlása valamilyen értelemben sztochasztikusan (átmenetenként, vagy elemi reakcióként) kiegyensúlyozott és milyen kapcsolat van ezen tulajdonságok fennállása és a determinisztikus részletes kiegyensúlyozottságnak a fennállása között (vö. még [44, 248—249. oldal]).

Ez a legutolsó kérdéscsoport *számolástechnikai szempontból* is fontos, ugyanis egy-egy kiegyensúlyozottsági feltétel az egyensúlyi eloszlás meghatározását jelentősen leegyszerűsíti. Ennek más alkalmazási területen — például tömegkiszolgáló rendszerek esetén is — van jelentősége [46].

FÜGGELÉK: Whittle egy tétele

Ismertetjük WHITTLE egy eredményét, amely bizonyos folyamatok stacionárius eloszlásának létezésére és alakjára vonatkozik.

F.1. DEFINÍCIÓ. Legyen $M \in \mathbb{N}$, és legyen $\xi: \Omega \times \mathbb{R}_0^+ \rightarrow \mathbb{N}_0^M$ egy olyan folyamat, amelyre

$$(F.1) \quad a_{j, j+e_m} = v_m,$$

$$(F.2) \quad a_{j, j-e_m} = \mu_m \varphi_m(j_m),$$

$$(F.3) \quad a_{j, j-e_m+e_n} = \lambda_{mn} \varphi_m(j_m),$$

$$a_{jl} = 0, \quad \text{ha } l-j \neq 0, \pm e_m, e_n - e_m$$

ahol

$$v_m, \mu_m, \lambda_{mn} \in \mathbb{R}_0^+; \quad j_m := \text{pr}_m j$$

és $\varphi_m: \mathbb{N}_0 \rightarrow \mathbb{R}_0^+$ olyan függvény, amelyre $\varphi_m(0)=0$; $m, n \in \{1, 2, \dots, M\}$ és $(\lambda_{mn})_{m, n \in \{1, 2, \dots, M\}}$ sztochasztikus mátrix. Az ilyen infinitézimális mátrixsal bíró folyamatokat egyszerű születési-halálozási folyamatoknak nevezzük.

F.1. TÉTEL [47]. A fenti egyszerű születési-halálozási folyamatnak akkor és csak akkor létezik egyetlen stacionárius eloszlása, ha a

$$(F.4) \quad \left(\mu_m + \sum_{n=1}^M \lambda_{mn} \right) \alpha_m - \sum_{n=1}^M \lambda_{nm} \alpha_n = v_m \quad (m \in \{1, 2, \dots, M\})$$

egyenletrendszernek (amely általában sokkal kevesebb egyenletből áll, mint a stacionárius eloszlásra vonatkozó!) létezik véges valós számokból álló $\{\alpha_m; m \in \{1, \dots, M\}\}$ megoldása, amely egyértelmű, és amelyre teljesül, hogy

$$\alpha_m < A_m \quad \text{konvergenciasugara} := \varrho(A_m), \quad (m \in \{1, \dots, M\}),$$

ahol

$$A_m(z) := \sum_{n \in \mathbb{N}_0} z^n / \varphi_m(1) \varphi_m(2) \dots \varphi_m(n) \\ (z \in \mathbb{C}, |z| < \varrho(A_m)),$$

és ekkor a $\{\pi_j; j \in \mathbb{N}_0^M\}$ stacionárius eloszlást a

$$(F.5) \quad \pi_j := \text{konst.} \prod_{m=1}^M \frac{\alpha_m^{j_m}}{[\varphi_m(1) \dots \varphi_m(j_m)]}$$

képlet (a „szorzatformula” egy speciális esete) adja meg. Az egyensúlyi eloszlásra teljesülnek a

$$(F.6) \quad \left[v_m + \sum_{n=1}^M \lambda_{nm} \varphi_n(j_{n+1}) T_n \right] T_m^{-1} \pi_j = \\ = \left[\mu_m + \sum_{n=1}^M \lambda_{mn} \right] \varphi_m(j_m) \pi_j \quad (m \in \{j, 2, \dots, M\}) \\ \sum_{n=1}^M \mu_n \varphi_n(j_n + 1) T_n \pi_j = \left[\sum_{n=1}^M v_n \right] \pi_j$$

összefüggések (amit úgy is mondhatunk a 3.3 definíció felhasználásával, hogy a folyamat sztochasztikusan átmenetenkénti részletesen kiegyensúlyozott). (Itt tetszőleges $f: N_0^M \rightarrow \mathbf{R}$, $m \in \{1, 2, \dots, M\}$ és $j \in N_0^M$ esetén

$$T_m f(j) := f(j + e_m).$$

Köszönetnyilvánítás. Köszönetünket fejezzük ki a lektoroknak és kollégáinknak a dolgozat megszerkesztésére és megfogalmazására vonatkozó megjegyzéseikért.

IRODALOM

- [1] ARNOLD, L., „A kémiai reakciók matematikai modelljeinek konzisztenciájáról”, *Alkalmazott Matematikai Lapok* (1981) 345—356.
- [2] BAILEY, N. T. J., *The Elements of Stochastic Processes with Applications to the Natural Sciences* (J. Wiley, New York—London—Sidney, 1964.)
- [3] BARTLETT, M. S., “Some evolutionary stochastic processes”, *J. Roy. Stat. Soc., Lond.* **B11** (1949) 211—223.
- [4] CHUNG, K. L., *Markov Chains with Stationary Transition Probabilities* (Springer, Berlin—Göttingen—Heidelberg, 1960.)
- [5] DELBRÜCK, M., “Statistical fluctuations in autocatalytic reactions”, *J. Chem. Phys.* **8** (1940) 120—124.
- [6] ÉRDI, P., *Az összetett kémiai reakció sztochasztikus modelljei és azok alkalmazásai* (Kandidátusi értekezés, Budapest, 1979.)
- [7] ÉRDI, P. and TÓTH, J., “Stochastic reaction kinetics = “Non-equilibrium thermodynamics” of the state space?”, *React. Kinet. Catal. Lett.* **4** (1976) 81—85.
- [8] ÉRDI, P. és TÓTH, J., „A kémiai reakció termodinamikájának sztochasztikus formulázásáról”, *A kémia újabb eredményei* **31** (1976) 177—298.
- [9] ÉRDI, P. and TÓTH, J., “Some comments on Prigogine’s theories”, *React. Kinet. Catal. Lett.* **11** (1979) 371—375.
- [10] ÉRDI, P., TÓTH, J., and HÁRS, V., “Some kinds of exotic phenomena in chemical systems”, in: *Qualitative Theory of Differential Equations* Ed. M. Farkas (János Bolyai Mathematical Society and North Holland Publishing Company, Amsterdam—London, 1981) 205—229.
- [10A] FELLER, W., *Bevezetés a valószínűségszámításba és alkalmazásaiba* (Műszaki Könyvkiadó, Budapest, 1978.)
- [11] FISHER, M. E., “The renormalization group in the theory of critical behavior” *Rev. Mod. Phys.* **46** (1974) 597—616.
- [12] FOX, R. F., “Gaussian stochastic processes in physics”, *Phys. Rep.* **48** (1978) 181—283.
- [13] GANS, P. J., “Open first-order processes”, *J. Chem. Phys.* **33** (1960) 691—694.
- [14] GARDINER, C. W. and CHATURVEDI, S., “The Poisson representation: A new technique for chemical master equations”, *J. Stat. Phys.* **17** (1977) 429—468.
- [15] GARDINER, C. W., MCNEIL, K. J., WALLS, D. F., and MATHESON, J. S., “Correlations in stochastic theories of chemical reactions”, *J. Stat. Phys.* **14** (1976) 309—331.
- [16] HANUSSE, P., *Contribution a l’étude des systemes dissipatifs chimiques. Simulation par une méthode de Monte Carlo. Etude d’une réaction oscillante*, (These, Université de Bordeaux I, 1976.)
- [17] INGARDEN, R. S. and KOSSAKOWSKI, A., “The Poisson probability distribution and information thermodynamics”, *Bull. Acad. Polon. Sci., Ser. math. astr. phys.* **19** (1971) 83—85.
- [18] IOSIFESCU, M. and TÁUTU, P., *Stochastic Processes and Applications in Biology and Medicine* Vol. 1, Theory (Ed. Acad., Bucureşti; Springer, Berlin—Heidelberg—New York, 1973.)
- [19] KAMPEN, N. G. VAN, “The equilibrium distribution of a chemical mixture”, *Phys. Lett.* **59A** (1976).
- [20] KAMPEN, N. G. VAN, “Fluctuations in closed and open non-linear systems”, *Physica* **83 B+C** (1976) 29—48.
- [21] KARLIN, S. and MCGREGOR, J., “The classification of birth and death processes”, *Trans. Amer. Math. Soc.* **86** (1957) 366—400.
- [22] KELLY, F. P., *Reversibility and Stochastic Networks* (J. Wiley, London, 1979).
- [23] KINGMAN, J. F. C., “Markov population processes”, *J. Appl. Prob.* **6** (1969) 1—18.
- [24] KURAMOTO, Y., “Fluctuations around steady states in chemical kinetics”, *Prog. Theor. Phys.* **49** (1973) 1782—1783.
- [25] KURAMOTO, Y., “Effects of diffusion on the fluctuations in open chemical systems”, *Prog. Theor. Phys.* **52** (1974) 711—713.

- [26] LANDAU, L. D., és LIFSIC, F. M., *Statistikus fizika* 1. (Elméleti fizika V.) (Tankönyvkiadó, Budapest, 1981).
- [27] LANDAUER, R., "Nonlinearity, multistability, fluctuations: reviewing the reviewers", *Am. J. Physiol.* **241** (1981) R107—R113.
- [28] LENGYEL, B., PRÉKOPA, A., RÉVÉSZ, P. und TÖRÖK, F., Über Kinetik und Gleichgewicht der Äquilibrations-Reaktion von linearen Methylpolysiloxanen II., *Z. Phys. Chem.* **208** (1957) 33—41.
- [29] LENGYEL, B., PRÉKOPA, A. und TÖRÖK, F., „Über Kinetik und Gleichgewicht der Äquilibrations-Reaktion von linearen Methylpolysiloxanen I., *Z. Phys. Chem.* **206** (1956) 161—168.
- [30] MALEK-MANSOUR, M. and NICOLIS, G., "A master equations description of local fluctuations", *J. Stat. Phys.* **13** (1975) 197—217.
- [31] MARTIN-LÖF, A., *Statistical Mechanics and the Foundations of Thermodynamics* (Lecture Notes in Physics 101, Springer, Berlin—Heidelberg—New York, 1979).
- [32] MEDGYESSY, P., „Sűrűségfüggvények és diszkrét eloszlások szuperpozícióinak felbontása”. *MTA III. Oszt. Közl.* **21** (1972) 129—200.
- [32A] MEDGYESSY P. és TAKÁCS L., *Valószínűségszámítás* (2. kiadás, Tankönyvkiadó, Budapest, 1966).
- [33] NICOLIS, G., ALLEN, P. and NYPELSEER, A. VAN, "Some remarks on the theory of fluctuations around nonequilibrium states", *Prog. Theor. Phys.* **52** (1974) 1481—1497.
- [34] NITZAN, A. and ROSS, J., "A comment on fluctuations around nonequilibrium steady states", *J. Stat. Phys.* **10** (1974) 379—390.
- [35] PRÉKOPA, A., „Hosszú láncmolekulák bomlási folyamatának valószínűségszámítási vizsgálata”, *MTA Alk. Mat. Int. Közl.* **2** (1953) 103—123.
- [36] PRÉKOPA, A., *Valószínűségelmélet műszaki alkalmazásokkal* (3. kiadás, Műszaki Könyvkiadó, Budapest, 1974).
- [37] PRIGOGINE, I., "Time, structure, and fluctuations", *Science* **201** (1978) 777—785.
- [38] RÉNYI, A., „Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével”, *MTA Alk. Mat. Int. Közl.* **2** (1953) 83—101.
- [39] RÉNYI, A., "On an extremal property of the Poisson process", *Annals Inst. Stat. Math.* **16** (1964) 129—133.
- [40] SIEGERT, A. J. F., "On the approach to statistical equilibrium", *Phys. Rev.* **76** (1940) 1708—1714.
- [41] TÓTH, J., "Poissonian stationary distribution in a class of detailed balanced reactions", *React. Kinet. Catal. Lett.* **18** (1981) 169—173.
- [42] TÓTH, J., „A formális reakciókinetika globális determinisztikus és sztochasztikus modelljéről és néhány alkalmazásáról”, *MTA SZTAKI Tanulmányok* **129** (1981).
- [43] TÓTH, J., „A kémiai reakciókinetika direkt és inverz feladatairól”, *Alkalmazott Matematikai Lapok* (1981) 253—269.
- [44] TÓTH, J. és ÉRDI, P., „A formális reakciókinetika modelljei, problémái és alkalmazásai”, *A kémia újabb eredményei* **41** (1978) 226—352.
- [45] TÓTH, J., and TÖRÖK, T. L., "Poissonian stationary distribution: A degenerate case of stochastic kinetics", *React. Kinet. Catal. Lett.* **13** (1980) 167—171.
- [46] TÖRÖK, T. L., "Balance equations in queueing networks" (Abstract), *Adv. Appl. Prob.* **10** (1978) 320.
- [47] WHITTLE, P., "Equilibrium distributions for an open migration process", *J. Appl. Prob.* **5** (1968) 567—571.
- [48] WILSON, K. G. and KOGUT, J., "The renormalization group, and the ϵ -expansion" *Physics Letters C. Physics Reports* **12C** (1974) 77—199.
- [49] ZUREK, W. H., and SCHIEVE, W. C., "Deviations from the Poisson behaviour of equilibrium fluctuations in a closed system", *J. Stat. Phys.* **22** (1980) 289—296.
- [50] Леонтович, М., «Основные уравнения кинетической теории газов с точки зрения теории случайных процессов», *Журн. эксперимент. и теорет. физики* **5** (1935) 211—231.

(Beérkezett: 1982. május 6.)

(Átdolgozva beérkezett: 1982. december 23.)

TÓTH JÁNOS
MTA SZTAKI
BUDAPEST, KENDE U. 13—17.
ÉRDI PÉTER
SOTE I. ANATÓMIAI INTÉZET, BUDAPEST
BUDAPEST, TÚZOLTÓ U. 58.
TÖRÖK LÁSZLÓ TURUL
MTA KFKI
BUDAPEST, KONKOLY THEGE M. ÚT 29—33.

SIGNIFICANCE OF THE POISSON DISTRIBUTION IN THE STOCHASTIC MODEL
OF COMPLEX CHEMICAL REACTIONS

J. TÓTH, P. ÉRDI, and T. L. TÖRÖK

The global stochastic model of complex chemical reactions is presented using *Vol'pert's definition* of the reactions. Evolutionary equations are written down for the model.

A sufficient condition is given for a reaction to have a unimodal stationary distribution. In the case of simple birth-and-death processes and of simple polynomial Markov population processes a necessary and sufficient condition is given for a reaction to have a *Poissonian stationary distribution*. This condition turns out to be fulfilled rather rarely—contrary to the expectation of part of the literature.

Connections between thermodynamic fluctuation theory and the character of the stationary distribution are sketched.

Several open mathematical problems are formulated that seem to be interesting from the standpoint of applications.

A basic statement by P. WHITTLE on *Markovian jump processes* is presented in an appendix.

KÁLMÁN-SZŰRŐK ASZIMPTOTIKUS VISELKEDÉSÉRŐL

HUHN EDIT

Szeged

Több dimenziós, részlegesen megfigyelhető Gauss-folyamatot tekintve a nem megfigyelhető komponensek optimális, lineáris előrejelzése hibájának aszimptotikus viselkedését vizsgáljuk.

1. Bevezetés

Tekintsünk egy $k+l$ dimenziós $(\mathfrak{Y}^T(t), \xi^T(t)) = ((\mathfrak{Y}_1(t), \dots, \mathfrak{Y}_k(t)), (\xi_1(t), \dots, \xi_l(t)))$, $0 \leq t \leq T$ Gauss-folyamatot, amelyik eleget tesz a következő sztochasztikus differenciálegyenlet-rendszernek:

$$(1.1) \quad \begin{cases} d\mathfrak{Y}(t) = [\mathfrak{a}_0(t) + a_1(t)\mathfrak{Y}(t) + a_2(t)\xi(t)]dt + \sum_{i=1}^2 b_i(t)d\mathbf{w}_i(t), \\ d\xi(t) = [A_0(t) + A_1(t)\mathfrak{Y}(t) + A_2(t)\xi(t)]dt + \sum_{i=1}^2 B_i(t)d\mathbf{w}_i(t), \end{cases}$$

ahol $\mathbf{w}_1^T = (w_1^1, \dots, w_k^1)$ és $\mathbf{w}_2^T = (w_1^2, \dots, w_l^2)$ független Wiener-folyamatok. A $(\mathfrak{Y}(0), \xi(0))$ kezdeti vektor független a $\mathbf{w}_1(t)$ és $\mathbf{w}_2(t)$ ($t \geq 0$) folyamatoktól.

Legyen $\mathbf{m}(t) = E(\mathfrak{Y}(t)|F_t^\xi)$ a feltételes várhatóérték-vektor ($F_t^\xi = \sigma\{\xi(s), 0 \leq s \leq t\}$) és $\gamma(t) = \{\gamma_{ij}(t)\}_{k \times k}$ a kovariancia-mátrix a

$$\gamma_{ij}(t) = E\{(\mathfrak{Y}_i(t) - m_i(t))(\mathfrak{Y}_j(t) - m_j(t))\}$$

elemekkel. Az $\mathbf{m}(t)$ nyilván F_t^ξ -mérhető becslése a $\mathfrak{Y}(t)$ -nek és optimális abban az értelemben, hogy

$$\text{Sp } \gamma(t) = \sum_{i=1}^k \gamma_{ii}(t) \leq \text{Sp } E\{(\mathfrak{Y}(t) - \mathbf{v}(t))(\mathfrak{Y}(t) - \mathbf{v}(t))^*\}$$

minden olyan F_t^ξ -mérhető $\mathbf{v}(t) = (v_1(t), \dots, v_k(t))^T$ vektor esetén, amelyre $\sum_{i=1}^k E|v_i(t)|^2 < \infty$ (Sp $\gamma(t)$ a becslés négyzetes középben vett hibája). A $(\mathfrak{Y}(t), \xi(t))$, $0 \leq t \leq T$ folyamat Gauss-volta miatt $\mathbf{m}(t)$ a $\xi(s)$, $s \leq t$ értékektől lineárisan függ.

Az (1.1) egyenletben szereplő $a_i(t)$, $A_i(t)$, $b_i(t)$, $B_i(t)$ együtthatókra tett bizonyos feltevések mellett érvényes a következő tétel [4]:

1.1. TÉTEL. Ha a $(\mathfrak{Y}(t), \xi(t))$ Gauss-folyamat eleget tesz az (1.1) egyenletrendszernek, akkor $m(t)$ és $\gamma(t)$ az

$$(1.2) \quad dm(t) = [a_0(t) + a_1(t)m(t) + a_2(t)\xi(t)]dt + [(b \circ B)(t) + \gamma(t)A_1^*(t)] \times \\ \times ((B \circ B)(t))^{-1} [d\xi(t) - A_0(t) + A_1(t)m(t) + A_2(t)\xi(t)dt],$$

$$(1.3) \quad \dot{\gamma}(t) = a_1(t)\gamma(t) + \gamma(t)a_1^*(t) - [(b \circ B)(t) + \gamma(t)A_1^*(t)] \times \\ \times ((B \circ B)(t))^{-1} [(b \circ B)(t) + \gamma(t)A_1^*(t)]^* + (b \circ b)(t)$$

egyenleteknek az $m(0) = E(\mathfrak{Y}(0)|\xi(0))$, $\gamma_{ij}(0) = E\{(\mathfrak{Y}_i(0) - m_i(0))(\mathfrak{Y}_j(0) - m_j(0))\}$ kezdeti feltételekhez tartozó megoldása.

A fenti tételben a

$$(b \circ b)(t) = b_1(t)b_1^*(t) + b_2(t)b_2^*(t)$$

$$(b \circ B)(t) = b_1(t)B_1^*(t) + b_2(t)B_2^*(t)$$

$$(B \circ B)(t) = B_1(t)B_1^*(t) + B_2(t)B_2^*(t)$$

jelöléseket használtuk.

Az (1.3) differenciálegyenlet *Riccati-típusú*, amelynek megoldását — általános esetben — nem tudjuk explicit alakban felírni. Ha az (1.3)-ban szereplő együtthatók t -től függetlenek, az egyenlet integrálható [2]. Az alábbiakban ilyen speciális esetekben fogjuk vizsgálni $\gamma(t)$ aszimptotikus viselkedését.

2. Folytonos, racionális spektrálsűrűségű folyamatok

Legyen $\zeta(t)$ reguláris Gauss-folyamat, amelynek spektrális sűrűségfüggvénye

$$f_\zeta(\lambda) = \frac{\left| \sum_{j=0}^q b_j(i\lambda)^{q-j} \right|^2}{\left| \sum_{j=0}^p a_j(i\lambda)^{p-j} \right|^2},$$

ahol $a_0 = 1$ és a $P(z) = z^p + a_1 z^{p-1} + \dots + a_p = 0$ egyenlet gyökeinek valós része negatív, $Q(z) = b_0 z^q + b_1 z^{q-1} + \dots + b_q$, $q < p$, $Q(z)$ és $P(z)$ együtthatói valósak, és nincs közös zéróhelyük. Az egyszerűség kedvéért tegyük fel, hogy $P(z)$ gyökei különbözők és legyenek ezek $\lambda_1, \dots, \lambda_p$. Ekkor létezik olyan $\zeta^T(t) = (\zeta^1(t), \dots, \zeta^p(t))$ stacionárius Gauss-folyamat, hogy

$$\zeta^1(t) = \zeta(t)$$

és $\zeta(t)$ komponensei a

$$(2.1) \quad \begin{cases} d\zeta^1(t) = \zeta^2(t)dt \\ \vdots \\ d\zeta^{p-q-1}(t) = \zeta^{p-q}(t)dt \\ d\zeta^{p-q}(t) = \sum_{i=1}^p d_i \zeta^i(t)dt + b_0 dw(t) \\ d\zeta^{p-q+1}(t) = \lambda_{p-q+1} \zeta^{p-q+1}(t)dt + dw(t) \\ \vdots \\ d\zeta^p(t) = \lambda_p \zeta^p(t)dt + dw(t) \end{cases}$$

lineáris sztochasztikus differenciálegyenleteknek tesznek eleget [1], ahol $w(t)$ a *standard Wiener-folyamat*, $E\{\zeta^j(0)w(t)\}=0$, $j=1, \dots, p$. Ha $q \geq 1$, akkor a $\zeta^1(t), \dots, \zeta^{p-q}(t)$ megfigyelhető, $\zeta^{p-q+1}(t), \dots, \zeta^p(t)$ pedig nem megfigyelhető komponensei a $\zeta(t)$ vektornak.

Legyen $\xi^T(t) = (\zeta^1(t), \dots, \zeta^{p-q}(t))$ és $\mathfrak{Y}^T(t) = (\zeta^{p-q+1}(t), \dots, \zeta^p(t))$. Ekkor az (1.1) egyenletek a következők:

$$(2.2) \quad \begin{cases} d\mathfrak{Y}(t) = a_1 \mathfrak{Y}(t) dt + b_2 dw(t) \\ d\xi(t) = [A_1 \mathfrak{Y}(t) + A_2 \xi(t)] dt + B_2 dw(t), \end{cases}$$

$$a_1 = \begin{bmatrix} \lambda_{p-q+1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_p \end{bmatrix}_{q \times q}, \quad b_2 = \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & & & \\ 0 & \dots & 0 & 1 \end{bmatrix}_{q \times (p-q)},$$

$$A_1 = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \\ d_{p-q+1} & \dots & d_p \end{bmatrix}_{(p-q) \times q}, \quad A_2 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & \dots & \dots & 0 & 1 \\ d_1 & d_2 & \dots & \dots & d_{p-q} \end{bmatrix}_{(p-q) \times (p-q)},$$

$$B_2 = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 1 \end{bmatrix}_{(p-q) \times (p-q)}, \quad w(t) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ w(t) \end{bmatrix}_{(p-q) \times 1}.$$

A $\gamma(t)$ -re vonatkozó (1.3) egyenlet pedig

$$(2.3) \quad \dot{\gamma}(t) = \gamma(t) \{a_1 - c_1 D_1^T\}^* + \{a_1 - c_1 D_1^T\} \gamma(t) - \gamma(t) D_1 D_1^T \gamma(t),$$

ahol

$$c_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{q \times 1}, \quad D_1 = \begin{bmatrix} d_{p-q+1} \\ \vdots \\ d_p \end{bmatrix}_{q \times 1}.$$

A (2.3) egyenlet megoldása

$$(2.4) \quad \gamma(t) = e^{(a_1 - c_1 D_1^T)t} \left\{ \gamma(0)^{-1} + \int_0^t e^{(a_1 - c_1 D_1^T)s} D_1 D_1^T e^{(a_1 - c_1 D_1^T)s} ds \right\}^{-1} e^{(a_1 - c_1 D_1^T)t}.$$

ahol $\gamma(0)$ egy, a $\zeta(t)$ folyamat által meghatározott mátrix-egyenlet megoldása [3].

A (2.4)-ben második tényezőként szereplő inverz mátrix — amint az az alábbiakból látható lesz — t elegendően nagy értékeire mindig létezik, ha $\gamma(0)$ invertálható.

Bizonyos kérdések vizsgálatánál fontos $\gamma(t)$ aszimptotikus viselkedésének ismerete. Például, ha az $(a_1, \dots, a_p, b_0, \dots, b_q)$ paraméterek aszimptotikusan effektív becsléseit keressük [1], vagy ha az a kérdés hogy a $\zeta(t)$ nem megfigyelhető komponensei mikor „aszimptotikusan megfigyelhetők”. Utóbbin azt értjük, hogy $\zeta(t)$ nem megfigyelhető komponensei négyzetes középben tetszőleges pontossággal becsülhetők a megfigyelhetők alapján, ha t elegendően nagy.

A következőkben elegendő feltételét adjuk annak, hogy $\lim_{t \rightarrow \infty} \gamma(t) = 0$ teljesüljön.

Legyen $\|\gamma(t)\| = \sup_{\|x\|=1} \|\gamma(t)x\|$ a szokásos operátor-norma, ahol x q -dimenziós vektor és $\|x\|$ az euklideszi norma. Nyilvánvaló, hogy $\lim_{t \rightarrow \infty} \gamma(t) = 0$ akkor és csakis akkor, ha $\lim_{t \rightarrow \infty} \|\gamma(t)\| = 0$.

2.1. TÉTEL. Ha az $a_1 - c_1 D_1^T$ mátrix sajátértékeinek valós része negatív, akkor $\lim_{t \rightarrow \infty} \gamma(t) = 0$.

Bizonyítás. A_1, \dots, A_q legyenek az $a_1 - c_1 D_1^T$ (nem szükségképpen különböző) sajátértékei. Az $e^{\{a_1 - c_1 D_1^T\}t}$ mátrix elemei $\sum_{j=1}^q P_{ik}^j(t) e^{A_j t}$ ($i, k = 1, \dots, q$) alakúak, ahol $P_{ik}^j(t)$ t -nek polinomja. Ezért létezik olyan $N \in \mathbb{R}$, hogy

$$(2.5) \quad \|e^{\{a_1 - c_1 D_1^T\}t}\| \leq N e^{-Kt}$$

$$\left(K = -\frac{1}{2} \max_{1 \leq i \leq q} \{\operatorname{Re} A_i\} > 0 \right).$$

Ennek felhasználásával kapjuk, hogy

$$(2.6) \quad \left\| \int_0^t e^{\{a_1 - c_1 D_1^T\}^* s} D_1 D_1^T e^{\{a_1 - c_1 D_1^T\}s} ds \right\| \leq \|D_1 D_1^T\| \times \\ \times \int_0^t \|e^{\{a_1 - c_1 D_1^T\}s}\|^2 ds \leq \|D_1 D_1^T\| \int_0^t N e^{-2Ks} ds \rightarrow 0 \quad (t \rightarrow \infty)$$

tehát

$$\left\| \left\{ \gamma(0)^{-1} + \int_0^t e^{\{a_1 - c_1 D_1^T\}^* s} D_1 D_1^T e^{\{a_1 - c_1 D_1^T\}s} ds \right\}^{-1} \right\|$$

korlátos. A fentiekből és

$$\|\gamma(t)\| \leq \|e^{\{a_1 - c_1 D_1^T\}t}\|^2 \left\| \left\{ \gamma(0)^{-1} + \int_0^t e^{\{a_1 - c_1 D_1^T\}^* s} D_1 D_1^T e^{\{a_1 - c_1 D_1^T\}s} ds \right\}^{-1} \right\|$$

alapján az állítás nyilvánvaló.

Megjegyzés. Az, hogy $P(z)$ gyökei negatív valós részűek, még nem biztosítja, hogy az $a_1 - c_1 D_1^T$ mátrix sajátértékei is ilyenek legyenek. További vizsgálatot igényel, hogy a $Q(z)$ polinomnak milyen tulajdonságúnak kell lennie ahhoz, hogy ez teljesüljön.

3. A Kálmán—Bucy-féle szűrő aszimptotikus tulajdonsága

Tekintsük a részlegesen megfigyelhető $(\mathfrak{Y}^T(t), \xi^T(t)) = ((y_1(t), \dots, y_k(t)), (\xi_1(t), \dots, \xi_l(t)))$, $t \geq 0$ Gauss-folyamatot, $(\xi(t))$ a megfigyelhető rész), amely megoldása a

$$(3.1) \quad d\mathfrak{Y}(t) = [a_1 \mathfrak{Y}(t) + a_2 \xi(t)] dt + b_1 d\mathbf{w}_1(t) + b_2 d\mathbf{w}_2(t) \\ d\xi(t) = [A_1 \mathfrak{Y}(t) + A_2 \xi(t)] dt + B_1 d\mathbf{w}_1(t) + B_2 d\mathbf{w}_2(t)$$

sztochasztikus differenciálegyenlet-rendszernek. Az a_i, A_i, b_i, B_i ($i=1, 2$) mátrixok t -től függetlenek és a $w_1(t), w_2(t)$ Wiener-folyamatok függetlenek és nem függenek a $(\mathfrak{g}(0), \xi(0))$ kezdeti vektortól.

Most az (1.3) egyenlet az

$$\begin{aligned}a &= a_1 - (b \circ B)(B \circ B)^{-1}A_1, \\b &= [(b \circ b) - (b \circ B)(B \circ B)^{-1}(b \circ B)^*]^{1/2}, \\B &= [B \circ B]^{1/2}, \quad A = A_1\end{aligned}$$

jelöléseket alkalmazva a következő [4]:

$$(3.2) \quad \dot{\gamma}(t) = a\gamma(t) + \gamma(t)a^* + bb^* - \gamma(t)A^*(BB^*)^{-1}A\gamma(t).$$

Nyilvánvaló, hogy ha $\lim_{t \rightarrow \infty} \gamma(t) = \gamma$ létezik, akkor kielégíti az

$$(3.3) \quad a\gamma + \gamma a^* + bb^* - \gamma A^*(BB^*)^{-1}A\gamma = 0$$

egyenletet.

3.1. TÉTEL. Ha a $D = a - \gamma A^*(BB^*)^{-1}A$ mátrix minden sajátértékének valós része negatív, akkor a (3.2) egyenlet minden $\gamma(t)$ megoldására $\lim_{t \rightarrow \infty} \gamma(t) = \gamma$.

Bizonyítás. Ha $\gamma(t)$ (3.2)-nek, γ pedig (3.3)-nak megoldása, akkor $\Gamma(t) = \gamma(t) - \gamma$ kielégíti a

$$(3.4) \quad \dot{\Gamma}(t) = D\Gamma(t) + \Gamma(t)D^* - \Gamma(t)A^*(BB^*)^{-1}A\Gamma(t)$$

egyenletet a $\Gamma(0) = \gamma(0) - \gamma$ kezdeti feltétellel. Nyilván $\lim_{t \rightarrow \infty} \gamma(t) = \gamma$ akkor és csakis akkor, ha $\lim_{t \rightarrow \infty} \Gamma(t) = 0$. A 2.1. tétel pedig pont azt állítja, hogy ha D sajátértékei negatív valós részűek, akkor $\lim_{t \rightarrow \infty} \Gamma(t) = 0$. Ezzel a tétel bizonyítását befejeztük.

Köszönetet mondok ARATÓ MÁTYÁS professzor úrnak, a probléma felvetéséért és értékes tanácsaiért.

IRODALOM

- [1] ARATÓ, M., "On sufficient statistics of Gaussian processes with rational spectral density function", *Anal. Math.* (megjelenés alatt).
- [2] REID, W. T., *Riccati Differential Equations* (Academic Press, New York and London, 1972).
- [3] ARATÓ, M., BENCZÜR, A., KRÁMLI, A., PERGEL, J., "Statistical problems of the elementary Gaussian processes" (MTA SZTAKI TANULMÁNYOK, 1975/41).
- [4] Липцер, Р. Ш., Ширяев, А. Н., *Статистика случайных процессов* (Наука, Москва, 1974).

(Beérkezett: 1982. november 10.)

HUHN EDIT

SZEGEDI ORVOSTUDOMÁNYI EGYETEM
SZÁMÍTÁSTECHNIKAI KÖZPONT
6720 SZEGED, PÉCSI U. 4/A.

ON THE ASYMPTOTIC BEHAVIOUR OF KALMAN FILTERS

E. HUHN

The multi-dimensional partially observable *Gaussian processes* are considered which are solutions of certain linear stochastic differential equation with constant coefficients. Using the fact that in this case the so called "filter equations" are integrable by quadratures the asymptotic behaviour of the error of prediction of the non-observable components is investigated.

A MAXIMUM LIKELIHOOD BECSLÉS ASZIMPTOTIKUS OPTIMALITÁSÁRÓL

PICK RÓBERT

Budapest

A dolgozatban a maximum likelihood módszer aszimptotikus tulajdonságaival foglalkozunk. A *Cramer-féle feltételrendszer* gyengítésével, nem feltétlenül azonos eloszlású változók esetén bizonyítjuk a módszer aszimptotikus konzisztens voltát és aszimptotikus normalitását. Ezen feltételek alapján gyakorlati szempontból is értékes támpontot kapunk a becslés egy értékeléséhez.

1. Bevezetés

A környező anyagi világ egyes eseményeiről, jelenségeiről sok esetben csak mérési eredmények útján juthatunk információkhoz. A megszerzett mérési eredmények általában nem biztosítják a közvetlen megismerést. Ezeknek valamely függvénye az, amely leírja a jelenségeket, eseményeket. A mérési eredmények a legtöbb esetben valószínűségi változókként foghatók fel és az ezeken értelmezett függvények — ha megfelelő tulajdonságúak — szintén valószínűségi változók lesznek, melyek ismeretlen paramétereinek becslése vezet a feladat megoldására. Általában nem okoz különösebb nehézséget egyetlen valószínűségi változón értelmezett egyváltozós függvények kezelése. Összetettebb a probléma, ha a mérési eredmények több dimenziós vektorváltozók és ezeken vektor-vektor vagy vektor-skalár függvényeket értelmezünk.

A nehézségek egyik oka — amely a gyakorlatban gyakran előfordul — az, hogy a mérési eredmények függetlensége ugyan biztosított, eloszlásuk ismert, de egy ezen értelmezett függvény alkalmazása során e kedvező tulajdonságok elvesznek és a számított értékek mint valószínűségi változók eloszlásának a meghatározása bonyolulttá válik. Így az ilyen jellegű feladatok megoldására általában a nem paraméteres módszerek használatosak. E feladatkörre egyszerű szemléletes példa lehet a háromszög csúcspontjának a meghatározása az alap és a rajtafekvő két szög ismeretében, mikor azoknak a hibaeloszlása $N(0, \sigma)$. A maximum likelihood becslés tulajdonságainak az alábbi feltételek mellett történő igazolása hasonló és más feladat megoldását is lehetővé teszi, paraméteres modellben, kihasználva a megfigyelhető valószínűségi változók eloszlásának ismeretét. A maximum likelihood becslések tulajdonságait az irodalomban vagy a *Wald*-, vagy a *Cramer-féle feltételrendszernek* megfelelően szokták igazolni. Míg az első általában a konzisztencia bizonyítására nyújt lehetőséget, a második feltételrendszer mellett aszimptotikus konzisztenciát és efficienciát lehetett eddig verifikálni. A következőkben a *Cramer-féle feltételrendszernek* megfelelő megszorítások mellett bizonyítjuk állításunkat, mivel ennek alapján a becslés kézenfekvő értékelését is közvetlenül előállíthatjuk, amelyet a gyakorlat számára is fontosnak tartunk.

2. Előkészületek

Vizsgálatainkat a következő feltételek mellett végeztük:

Legyen $\xi_1, \xi_2, \xi_3, \dots$ független valószínűségi változók végtelen sorozata, és tegyük fel, hogy ξ_k sűrűségfüggvénye $f(x, \vartheta_{01} \dots \vartheta_{0r}, \tau_{k1}, \dots, \tau_{kl})$, ahol $\vartheta_0 = (\vartheta_{01}, \dots, \vartheta_{0r})$ ismeretlen r -dimenziós vektor, a $\{\tau_k = (\tau_{k1}, \dots, \tau_{kl}); k=1, 2, \dots\}$ sorozat pedig ismert l -dimenziós vektorokból áll.

A ϑ_0 ismeretlen paraméterről csak annyit tudunk, hogy az r -dimenziós E^r euklideszi tér θ részhalmazának eleme, és van olyan $\omega \subset \theta$ környezete, amelyben teljesülnek a következő feltételek:

A) Az $A_k = \{x: f(x, \vartheta, \tau_k) > 0\}$ halmaz rögzített k mellett független ϑ -tól, a

$$g(x, \vartheta, \tau_k) = \log F(x, \vartheta, \tau_k)$$

függvény ϑ szerint háromszor parciálisan deriválható, a deriváltak folytonosak ϑ -ban minden $\vartheta \in \omega$ és majdnem minden $x \in A_k$ mellett. (A továbbiakban felteszszük, hogy ha $x \notin A_k$, akkor $g(x, \vartheta, \tau_k) = 0$.)

B) Minden i, j, l -re $i, j, l = 1, 2, \dots, r$.

$$\int_{A_k} \sup_{\vartheta' \in \omega} |f_{\vartheta_i}(x, \vartheta', \tau_k)| dx < +\infty$$

$$\int_{A_k} \sup_{\vartheta' \in \omega} |f_{\vartheta_i \vartheta_j}(x, \vartheta', \tau_k)| dx < +\infty$$

valamint

$$\int_{A_k} \left(\left| \frac{\partial}{\partial \vartheta_i} \log f(x, \vartheta, \tau_k) \right|^2 \right) f(x, \vartheta, \tau_k) dx < +\infty$$

teljesül minden k -ra, továbbá minden k -hoz van olyan γ_k véletlen szám, melyre $M(\gamma_k) < +\infty$, és

$$\sup_{\vartheta' \in \omega} \left| \frac{\partial^3 \log f(\xi_k, \vartheta', \tau_k)}{\partial \vartheta_i \partial \vartheta_j \partial \vartheta_l} \right| < \gamma_k$$

teljesül.

C/ 1. Létezik a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n M \left(\frac{\partial^2 g(\xi_k, \vartheta_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \right) = L_{i,j}$$

határérték minden $1 \leq i, j \leq r$ mellett, és az $L = (L_{i,j})_{i,j=1}^r$ mátrix reguláris.

C/ 2. Az $M \left[\left(\frac{\partial^2 g(\xi_k, \vartheta_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \right)^2 \right]$ sorozat k -ban egyenletesen korlátos minden $1 \leq i, j \leq r$ mellett.

C/ 3. Az $q_k = \frac{\partial g(\xi_k, \vartheta_0, \tau_k)}{\partial \vartheta}$ valószínűségi vektorváltozók eloszlására teljesül, hogy tetszőleges $\varepsilon > 0$ mellett

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \int_{\|x\| > \varepsilon \sqrt{n}} \|x\|^2 dG_k = 0.$$

Szükségünk lesz a következő két segédételre:

2.1. TÉTEL (RÉNYI [22] 313. o.).

Legyenek ξ_k ($k=1, 2, \dots$) független valószínűségi változók. Tegyük fel, hogy $M_k = M(\xi_k)$ és $D_k = D(\xi_k)$, $k=1, 2, \dots$ végesek. Tegyük fel továbbá, hogy teljesülnek a következő feltételek:

a) létezik a $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n M_k = M$ véges határérték;

b) fennáll, hogy $\lim_{n \rightarrow \infty} \frac{S_n}{n} = 0$, ahol $S_n = \sqrt{\sum_{k=1}^n D_k^2}$, (amihez elég, ha a D_k sorozat egyenletesen korlátos).

Ekkor a $\frac{1}{n} \sum_{k=1}^n \xi_k$ valószínűségi változó sztochasztikusan tart M -hez.

2.2. TÉTEL (CRAMER [6] 114. o.).

Legyen a $\mathbf{q}_1, \mathbf{q}_2, \dots$ r -dimenziós független valószínűségi változók sorozata olyan, hogy minden \mathbf{q}_n G_n eloszlásfüggvényére teljesüljön, hogy az első momentumai eltűnnek és a $\mu_{i,j}^{(k)}$ második momentumok végesek.

Tegyük fel, hogy $n \rightarrow \infty$ -re a következő két feltétel teljesül:

$$\frac{1}{n} \sum_{k=1}^n \mu_{i,j}^{(k)} \rightarrow \mu_{i,j} \quad i, j = 1, 2, \dots, r,$$

ahol a $\mu_{i,j}$ számok nem mind 0-ák és

$$\frac{1}{n} \sum_{k=1}^n \int_{\|\mathbf{x}\| > \varepsilon \sqrt{n}} \|\mathbf{x}\|^2 dG_k \rightarrow 0$$

minden $\varepsilon > 0$ -ra, ahol $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_r^2}$. Akkor a $(\mathbf{q}_1 + \mathbf{q}_2 + \dots + \mathbf{q}_n) \frac{1}{\sqrt{n}}$ változó eloszlásfüggvénye konvergál ahhoz a normális eloszláshoz, melynek az első momentuma nulla és a második momentumai a $\mu_{i,j} - k$.

3. A likelihood tétel

Tegyük fel, hogy a $\xi_1, \dots, \xi_k \dots$ sorozatra teljesülnek az előző pont A, B, C feltételei.

3.1. TÉTEL.

Legyen

$$L(\mathfrak{g}) = \sum_{k=1}^n g(\xi_k, \mathfrak{g}, \tau_k)$$

az ún. likelihood függvény, és jelöljük L lokális maximumainak a helyét H_n -nel. Tegyük fel, hogy az f sűrűségfüggvényre, és a $\{\tau_k, k=1, 2, \dots\}$ sorozatra teljesülnek az A, B, C feltételek. Akkor van olyan β_n sorozat, amelynek eloszlása tart az $N(0, \mathbf{D}^{-1})$ eloszláshoz, ahol $\mathbf{D} = -\mathbf{L}$, és tetszőleges $\eta > 0$ és $\alpha > 0$ számokhoz van olyan n_0 , hogy ha $n > n_0$, akkor

$$P(\exists \mathfrak{g}_n \in H_n, \text{ melyre } \|\sqrt{n}(\mathfrak{g}_n - \mathfrak{g}_0) - \beta_n\| < \alpha) \cong 1 - \eta.$$

Bizonyítás. Fejtsük *Taylor-sorba* az $L(\mathfrak{g}) = \sum_{k=1}^n g(\xi_k, \mathfrak{g}, \tau_k)$ likelihood függvényt az igazi paraméter körül:

$$L(\mathfrak{g}) = L(\mathfrak{g}_0) + L'(\mathfrak{g}_0)^T(\mathfrak{g} - \mathfrak{g}_0) + \frac{1}{2}(\mathfrak{g} - \mathfrak{g}_0)^T L''(\mathfrak{g}_0)(\mathfrak{g} - \mathfrak{g}_0) + R(\mathfrak{g}).$$

Belátjuk, hogy $\frac{1}{n} L''(\mathfrak{g}_0)$ sztochasztikusan tart az \mathbf{L} mátrixhoz, és az negatív definit.

Tekintsük az $L''(\mathfrak{g}_0)$ mátrix (i, j) -edik elemét, amely

$$\sum_{k=1}^n \frac{\partial^2 g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j}$$

alakú.

Mivel $M \left[\left(\frac{\partial^2 \log f(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \right)^2 \right]$ a $C/2$. feltétel szerint egyenletesen korlátos és a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum M \left(\frac{\partial^2 g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \right) = L_{i,j}$$

határérték létezik, alkalmazható a 2.1. tétel a $\xi_k = \frac{\partial^2 g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j}$ változóra, rögzített $1 \leq i, j \leq r$ mellett.

A tétel feltételei feljesülnek a $C/1$. és $C/2$. feltételek miatt. Kapjuk, hogy

$$\frac{1}{n} \sum_{k=1}^n \frac{\partial^2 g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \rightarrow (L_{i,j}),$$

tehát $\frac{1}{n} L''(\mathfrak{g}_0) \rightarrow \mathbf{L}$ -hez. Így minden $\varepsilon > 0$ és $\eta > 0$ -hoz létezik olyan N , hogy ha $n > N$, akkor

$$P \left(\left\| \frac{1}{n} L''(\mathfrak{g}_0) - \mathbf{L} \right\| > \varepsilon \right) < \eta.$$

(A továbbiakban, ha mást nem mondunk $\| \cdot \|$ az euklideszi normát jelenti.)

Vizsgáljuk továbbá az

$$\frac{1}{n} \sum_{k=1}^n M \left(\frac{\partial^2 g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j} \right)$$

összeget.

Ez az integrálás és differenciálás felcserélhetősége miatt

$$\frac{1}{n} \sum_{k=1}^n M \left(- \frac{\partial g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i} \cdot \frac{\partial g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_j} \right)$$

alakba írható.

Az összegben szereplő tagok pedig nem mások, mint a $\mathbf{Q}_k = \frac{\partial g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \mathfrak{g}}$ valószínűségi változók kovariancia-mátrixának az elemei, ami pozitív szemidefinit. Mivel \mathbf{L} -ről feltételeztük, hogy reguláris, így a fentiek szerint \mathbf{L} csak negatív definit lehet.

Tekintsük az

$$L(\mathfrak{g}) = L(\mathfrak{g}_0) + \frac{1}{\sqrt{n}} L'(\mathfrak{g}_0)^T (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T L(\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + \\ + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + R(\mathfrak{g})$$

felbontást.

Definiáljuk az

$$\tilde{L}(\mathfrak{g}) = L(\mathfrak{g}_0) + \frac{1}{\sqrt{n}} L'(\mathfrak{g}_0)^T (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T L(\mathfrak{g} - \mathfrak{g}_0) \sqrt{n}$$

függvényt.

Szükségünk lesz a következő lemmára.

3.1. LEMMA. Legyen $\xi_1, \xi_2, \dots, \xi_n$ független valószínűségi változók egy sorozata, melyeknek rendre $f(x_1, \mathfrak{g}_0, \tau_1), f(x_2, \mathfrak{g}_0, \tau_2), \dots, f(x_n, \mathfrak{g}_0, \tau_n)$ a sűrűségfüggvényei és kielégítik a tétel feltételeit.

Legyen továbbá

$$\frac{\partial L(\mathfrak{g}_0)}{\partial \mathfrak{g}} = \frac{\partial}{\partial \mathfrak{g}} \left(\sum_{k=1}^n g(\xi_k, \mathfrak{g}, \tau_k) \right)_{\mathfrak{g}=\mathfrak{g}_0}$$

Ekkor a $\beta_n = -\frac{1}{\sqrt{n}} L^{-1} \left(\frac{\partial L(\mathfrak{g}_0)}{\partial \mathfrak{g}} \right)$ valószínűségi vektorváltozó eloszlásfüggvényére $\lim_{n \rightarrow \infty} F_n(\mathbf{x}) = F(\mathbf{x})$, ahol $F(\mathbf{x})$ egy $N(0, D^{-1})$ eloszlású valószínűségi vektorváltozó eloszlásfüggvénye ($D = -L$), valamint rögzített $\eta > 0$ -hoz van olyan $C > 0$ véges szám, hogy minden n -re teljesül

$$P(\|\beta_n\| < C) \cong 1 - \eta.$$

ahol

$$\|\beta_n\| = \max_{1 \leq i \leq r} |(\beta_n)_i|.$$

Bizonyítás. Mindenekelőtt azt kell belátni, hogy β_n valószínűségi vektorváltozó $F_n(\mathbf{x})$ eloszlás függvényének határeloszlása egy $N(0, D^{-1})$ eloszlás. Ezen állítás igazolásához felhasználjuk a 2.2. tételt.

Tekintsük az r -dimenziós valószínűségi változók alábbi sorozatát:

$$\mathbf{q}_k = \frac{\partial g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \mathfrak{g}}, \quad k = 1, 2, \dots$$

A ξ_k -k független valószínűségi változók voltak, így az r -dimenziós \mathbf{q}_k valószínűségi vektorváltozók sorozata is független.

A második momentumok mátrixának létezése a tétel C/2. feltétele alapján következik és a

$$M \left(\frac{\partial g(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \mathfrak{g}} \right) = 0$$

feltétel az integrálás és differenciálás felcserélhetősége miatt teljesül, mivel

$$\int_{A_k} \frac{\partial g(x, \mathfrak{g}_0, \tau_k)}{\partial \mathfrak{g}} dF_k(x) = \int_{A_k} \frac{1}{f(x, \mathfrak{g}_0, \tau_k)} f_{\mathfrak{g}}(x, \mathfrak{g}_0, \tau_k) f(x, \mathfrak{g}_0, \tau_k) dx,$$

de $\int f dx = 1$ miatt

$$\frac{\partial}{\partial \mathfrak{g}} \int_{A_k} f dx = \int_{A_k} f_{\mathfrak{g}} dx = 0.$$

A 2.2. tétel feltételei tételünk C) feltételeiből következnek.

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \mathbf{q}_k \text{ eloszlásfüggvénye konvergál ahhoz az } r\text{-dimenziós normális eloszláshoz,}$$

melynek várható értéke $\mathbf{0}$, kovariancia mátrixa $-\mathbf{L}$. Legyen $\beta_n = -\mathbf{L}^{-1} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \mathbf{q}_k \right)$.

Ekkor $\beta_n F_n(\mathbf{x})$ eloszlása tart egy normális eloszláshoz, melynek várható értéke $\mathbf{0}$. A kovariancia affin transzformáció hatására tenzorként transzformálódik (RÉNYI [7] 209. o.). Figyelembe véve, hogy $(-\mathbf{L}^{-1})^T = -\mathbf{L}^{-1}$, a kovariancia $(-\mathbf{L}^{-1})(-\mathbf{L})(-\mathbf{L}^{-1}) = -\mathbf{L}^{-1} = \mathbf{D}^{-1}$. Tehát

$$\beta_n \xrightarrow{L} \beta$$

Azt kell még megmutatnunk, hogy van olyan $C > 0$ véges szám, hogy minden n -re teljesül

$$P(\|\beta_n\| < C) \geq 1 - \eta$$

Legyen $1 > \eta > 0$ tetszőleges, és $\|\beta\| = \sup_{1 \leq i \leq r} |\beta_i|$; $\|\beta_n\| = \sup_{1 \leq i \leq r} |\beta_{n_i}|$.

Mivel $\|\beta\| > 0$, $M(\|\beta\|) < +\infty$, így a Markov-egyenlőtlenség miatt $C(\beta, \eta)$ választható úgy, hogy teljesül

$$P(\|\beta\| < C(\beta, \eta)) \geq 1 - \frac{\eta}{2},$$

mert

$$P(\|\beta\| \geq \lambda M(\|\beta\|)) \leq \frac{1}{\lambda},$$

ahol

$$\lambda = \frac{1}{\eta/2} \quad \text{és} \quad P\left(\|\beta\| \geq \frac{M(\|\beta\|)}{\eta/2}\right) \leq \frac{\eta}{2}.$$

A komplementer eseményre térve, a

$$P\left(\|\beta\| < \frac{M(\|\beta\|)}{\eta/2}\right) \geq 1 - \frac{\eta}{2}$$

egyenlőtlenségből következik, hogy

$$C(\beta, \lambda) = \frac{M(\|\beta\|)}{\eta/2}$$

választással teljesül a

$$P(\|\beta\| < C(\beta, \eta)) \geq 1 - \frac{\eta}{2}$$

egyenlőtlenség.

Feltehető, hogy $C(\beta, \eta)$ olyan, hogy

$$F_n(C(\beta, \eta)) \rightarrow F(C(\beta, \eta))$$

Az $F_n(x) \rightarrow F(x)$ -ből következik, hogy adott $\eta > 0$ mellett választható n_0 , hogy minden $n > n_0$ -ra és $C(\beta, \eta)$ -ra teljesül,

$$P(|(\|\beta_n\| < C(\beta, \eta)) - P(\|\beta\| < C(\beta, \eta))| < \eta/2$$

és

$$P(\|\beta\| < C(\beta, \eta)) \geq 1 - \eta/2.$$

Alkalmazzuk újra a *Markov-egyenlőtlenséget* minden n -re, melyre $1 \leq n \leq n_0$

$$P\left(\|\beta_n\| < \frac{M(\|\beta_n\|)}{\eta/2}\right) \geq 1 - \eta/2.$$

Ekkor $C(n, \eta) = \frac{M(\|\beta_n\|)}{\eta/2}$ minden n -re $1 \leq n \leq n_0$ és rögzített $\eta > 0$ -ra előáll.

Ha

$$C = \max(C(\beta, \eta); \max_{1 \leq n \leq n_0} C(n, \eta))\text{-nak}$$

választjuk, akkor nyilván minden n -re kielégíti a

$$P(\|\beta_n\| < C) \geq 1 - \eta$$

egyenlőtlenséget, és ezzel a lemmát bebizonyítottuk.

Most megmutatjuk, hogy $\tilde{L}(\mathfrak{g})$ felveszi abszolút maximumát egy olyan $\tilde{\mathfrak{g}}_n$ helyen, melyre $\sqrt{n}(\tilde{\mathfrak{g}}_n - \mathfrak{g}_0) = \beta_n$.

$\tilde{L}(\mathfrak{g})$ átírható az alábbi alakba:

$$\tilde{L}(\mathfrak{g}) = L(\mathfrak{g}_0) + \frac{1}{2} \sqrt{n} \left(\mathfrak{g} - \mathfrak{g}_0 - \frac{\beta_n}{\sqrt{n}} \right)^T \mathbf{L} \left(\mathfrak{g} - \mathfrak{g}_0 - \frac{\beta_n}{\sqrt{n}} \right) \sqrt{n} - \frac{1}{2} \beta_n^T \mathbf{L} \beta_n = N.$$

Ugyanis

$$\begin{aligned} N &= L(\mathfrak{g}_0) + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \mathbf{L} (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} - \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \mathbf{L} \frac{\beta_n}{\sqrt{n}} \sqrt{n} - \\ &\quad - \frac{1}{2} \sqrt{n} \frac{\beta_n^T}{\sqrt{n}} \mathbf{L} (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n}. \end{aligned}$$

Felhasználva, hogy $\beta_n = -\mathbf{L}^{-1} \left(\frac{L'(\mathfrak{g}_0)}{\sqrt{n}} \right)$ kapjuk:

$$\begin{aligned} N &= L(\mathfrak{g}_0) + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \mathbf{L} (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + \frac{1}{2} \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \mathbf{L} \mathbf{L}^{-1} \left(\frac{L'(\mathfrak{g}_0)}{n} \right) \sqrt{n} + \\ &\quad + \frac{1}{2} \sqrt{n} \frac{L'(\mathfrak{g}_0)^T}{n} \mathbf{L}^{-1} \mathbf{L} (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} = \tilde{L}(\mathfrak{g}). \end{aligned}$$

Vegyük észre, hogy

$$K = L(\mathfrak{g}_0) - \frac{1}{2} \beta_n^T L \beta_n$$

független \mathfrak{g} -től.

Az $u(x) = x^T L x$ kvadratikus alaknak az $x=0$ helyen (itt r -dimenziós vektorok szerepelnek) abszolút maximuma van, hiszen L negatív definit.

Tehát $\tilde{L}(\mathfrak{g})$ a maximumát a $\tilde{\mathfrak{g}}_n = \mathfrak{g}_0 + \frac{\beta_n}{\sqrt{n}}$ helyen veszi fel, a maximum értéke K .

Tekintsük az

$$L(\mathfrak{g}) = \tilde{L}(\mathfrak{g}_0) + \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + R(\mathfrak{g})$$

függvényt, ahol

$$R(\mathfrak{g}) = \frac{1}{\sqrt{n}} \left(\frac{1}{6n} \sum_{i=1}^r \sum_{j=1}^r \sum_{l=1}^r \frac{\partial^3 L(\mathfrak{g}^*)}{\partial \mathfrak{g}_i \partial \mathfrak{g}_j \partial \mathfrak{g}_l} \sqrt{n} (\mathfrak{g}_i - \mathfrak{g}_{0i}) \sqrt{n} (\mathfrak{g}_j - \mathfrak{g}_{0j}) \sqrt{n} (\mathfrak{g}_l - \mathfrak{g}_{0l}) \right).$$

Az $\alpha > 0$ adott számhoz válasszuk n -t olyan nagyra, hogy

$$K = \{\|u\| \leq \alpha; \text{ ahol } u = \sqrt{n} (\mathfrak{g} - \tilde{\mathfrak{g}}_n)\}$$

környezet már $\|\beta_n\| < C$ mellett θ része legyen és válasszuk ehhez $\varepsilon > 0$ számot, hogy teljesüljön

$$\max_{\mathfrak{g} \in \max K} \tilde{L}(\mathfrak{g}) = \max_{\|u\| < \alpha} u^T L u + K < K - \varepsilon.$$

Ezek a mennyiségek $\tilde{L}(\mathfrak{g})$ szélsőérték tulajdonsága miatt léteznek.

Azt akarjuk megmutatni, hogy alkalmas n_0 választással elérhető, hogy

$$\left| \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0) + R(\mathfrak{g}) \right| < \frac{\varepsilon}{3}$$

teljesül nagy valószínűséggel, ha $n > n_0$. Válasszuk meg n_1 -et úgy, hogy $n > n_1$ mellett

$$P \left(\left\| \frac{1}{n} L''(\mathfrak{g}_0) - L \right\| < \frac{\varepsilon}{6(C+\alpha)^2} \right) > 1 - \frac{\eta}{6}$$

teljesüljön.

Ez $\frac{1}{n} L''(\mathfrak{g}_0)$ sztochasztikus konvergenciája miatt lehetséges.

Ekkor minden $\mathfrak{g} \in K$ mellett $\sqrt{n} \|\mathfrak{g} - \mathfrak{g}_0\| \leq C_2 = C + \alpha$ tehát,

$$\left| \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} \right| \leq \frac{C_2^2 \varepsilon}{6C_2^2} \leq \frac{\varepsilon}{6}.$$

Vizsgáljuk $R(\mathfrak{g})$ -t.

$$\begin{aligned} |R(\mathfrak{g})| &= \frac{1}{\sqrt{n}} \left| \frac{1}{6n} \sum_{i=1}^r \sum_{j=1}^r \sum_{l=1}^r \frac{\partial^3 L(\mathfrak{g}^*)}{\partial \vartheta_i \partial \vartheta_j \partial \vartheta_l} \sqrt{n} (\vartheta_i - \vartheta_{0i}) \sqrt{n} (\vartheta_j - \vartheta_{0j}) \sqrt{n} (\vartheta_l - \vartheta_{0l}) \right| \leq \\ &\leq \frac{1}{\sqrt{n}} \frac{1}{6n} \sum_{i=1}^r \sum_{j=1}^r \sum_{l=1}^r \sup \left| \frac{\partial^3 L(\mathfrak{g}^*)}{\partial \vartheta_i \partial \vartheta_j \partial \vartheta_l} \right| \sqrt{n} (\vartheta_i - \vartheta_{0i}) \sqrt{n} (\vartheta_j - \vartheta_{0j}) \sqrt{n} (\vartheta_l - \vartheta_{0l}). \end{aligned}$$

Mivel a tétel feltételei szerint minden k -ra létezik

$$\gamma_k \equiv \sup_{\mathfrak{g}' \in \omega} \left| \frac{\partial^3 \log f(\xi_k, \mathfrak{g}_0, \tau_k)}{\partial \vartheta_i \partial \vartheta_j \partial \vartheta_l} \right|,$$

így a *Markov-egyenlőtlenség* miatt van olyan C_1 , amelyre n -től függetlenül:

$$P\left(\frac{1}{n} \sum_{k=1}^n \gamma_k < C_1\right) \equiv 1 - \frac{\eta}{6}.$$

Az eseménytérnek azon a részén, ahol a szóban forgó esemény teljesül, érvényes a következő becslés:

$$R(\mathfrak{g}) \leq \frac{1}{6\sqrt{n}} \left| \sum_{i=1}^r \sum_{j=1}^r \sum_{l=1}^r C_1^3 C_2^3 \right| \leq \frac{1}{6\sqrt{n}} r^3 C_1^3 C_2^3 < \frac{\varepsilon}{6},$$

ha

$$n_2 > n = \left(\frac{C_1^3 C_2^3 r^3}{\varepsilon} \right)^2.$$

Ha $n > n_0$, akkor igaz, hogy

$$\begin{aligned} &\left| \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} \mathbf{L}''(\mathfrak{g}_0) - \mathbf{L} \right) (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} + R(\mathfrak{g}) \right| \leq \\ &\leq \left| \sqrt{n} (\mathfrak{g} - \mathfrak{g}_0)^T \left(\frac{1}{n} \mathbf{L}''(\mathfrak{g}_0) - \mathbf{L} \right) (\mathfrak{g} - \mathfrak{g}_0) \sqrt{n} \right| + |R(\mathfrak{g})| \leq \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3} \end{aligned}$$

minden $\mathfrak{g} \in K$ -ra.

Ekkor a választott K nyílt környezetre egyrészt igaz, hogy

$$\inf \tilde{L}(\mathfrak{g}) > K - \varepsilon,$$

és

$$L(\tilde{\mathfrak{g}}_n) = \tilde{L}(\mathfrak{g}_n) + (L(\tilde{\mathfrak{g}}_n) - \tilde{L}(\tilde{\mathfrak{g}}_n)) > K - \frac{\varepsilon}{3};$$

másrészt a határon

$$\max L(\mathfrak{g}) \leq K - \frac{2\varepsilon}{3}.$$

Tehát az $L(\mathfrak{g})$ függvény valamely $\mathfrak{g} \in K$ helyen a *Weierstrass-tétel* miatt felveszi a maximumát, feltéve, hogy

$$\|\beta_n\| < C; \left\| \frac{1}{n} L''(\mathfrak{g}_0) - L \right\| < \frac{\varepsilon}{6C_2^2}; \quad \frac{1}{n} \sum_{k=1}^n \gamma_k < C_1.$$

Annak viszont, hogy ezek teljesülnek, nagy a valószínűsége, mert

$$\begin{aligned} P\left(\|\beta_n\| < C; \left\| \frac{1}{n} L''(\mathfrak{g}_0) - L \right\| < \frac{\varepsilon}{6C_2^2}; \quad \frac{1}{n} \sum_{k=1}^n \gamma_k < C_1\right) &\equiv \\ &\equiv 1 - P(\|\beta_n\| > C) + P\left(\left\| \frac{1}{n} L''(\mathfrak{g}_0) - L \right\| > \frac{\varepsilon}{6C_2^2}\right) + P\left(\frac{1}{n} \sum_{k=1}^n \gamma_k > C_1\right) > 1 - \eta. \end{aligned}$$

Így a bizonyítást befejeztük.

A következő tétel az unicitás problémájával foglalkozik.

3.2. TÉTEL. Tegyük fel, hogy teljesülnek a 3.1. tétel feltételei. Ekkor minden $\eta > 0$ -hoz létezik \mathfrak{g}_0 körüli ω környezet és n_0 olyan, hogy ha $n > n_0$, akkor

$$P(\exists \mathfrak{g}_1 \neq \mathfrak{g}_2; \mathfrak{g}_1, \mathfrak{g}_2 \in \omega \text{ és } L(\mathfrak{g})\text{-nak } \mathfrak{g}_1, \mathfrak{g}_2\text{-ben lok. max. van}) < \eta.$$

Bizonyítás. Tekintsük $L(\mathfrak{g})$ likelihood függvényt és $L''(\mathfrak{g})$ -t fejtsük *Taylor-sorba* \mathfrak{g}_0 valódi paraméter körül.

$$\frac{1}{n} L''(\mathfrak{g}) = \frac{1}{n} L''(\mathfrak{g}_0) + (\mathfrak{g} - \mathfrak{g}_0) \frac{L'''(\mathfrak{g}^*)}{n}.$$

A 3.1. tételből tudjuk egyrészt, hogy $\frac{1}{n} L''(\mathfrak{g}_0)$ tart sztochasztikusan L -hez és minden $\varepsilon > 0$ -hoz létezik n_0 olyan, hogy ha $n > n_0$, akkor

$$P\left(\left\| \frac{1}{n} L''(\mathfrak{g}_0) - L \right\| > \varepsilon\right) < \eta/2,$$

másrészt pedig $L'''(\mathfrak{g})$ sztochasztikusan egyenletesen korlátos. Így létezik olyan K , hogy

$$P\left(\left\| \frac{L'''(\mathfrak{g}^*)}{n} \right\| > K\right) < \eta/2.$$

Ekkor

$$\frac{1}{n} L''(\mathfrak{g}) = L + \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) + (\mathfrak{g} - \mathfrak{g}_0) \frac{L'''(\mathfrak{g}^*)}{n} \quad \text{és}$$

$$\left\| \frac{1}{n} L''(\mathfrak{g}) - L \right\| \leq \varepsilon + \|\mathfrak{g} - \mathfrak{g}_0\| \cdot K$$

teljesül egy legalább $1 - \eta$ mértékű halmazon.

Ismeretes, hogy egy szimmetrikus mátrix pontosan akkor pozitív definit, ha a főminorai pozitívak (*Sylvester-kritérium*). Mivel a minorok a mátrix elemeinek

folytonos függvényei (velük képzett determinánsok), ha $-\frac{1}{n}L''(\mathfrak{g})$ elég közel van $-L$ -hez, akkor $-\frac{1}{n}L'(\mathfrak{g})$ főminorai is pozitívak. Ez elérhető, ha ε és $\|\mathfrak{g}-\mathfrak{g}_0\|$ kicsi. $\varepsilon > 0$ tetszőlegesen választható, és utóbbi is elérhető az ω környezet alkalmas választásával és a $\mathfrak{g} \in \omega$ feltevessel.

4. Összegezés

Ebben a pontban összefoglaljuk és kiegészítjük a 3. pont eredményeit.

4.1. TÉTEL. Tegyük fel, hogy teljesülnek a 3.1. tétel feltételei. Ekkor létezik olyan \mathfrak{g}_n sorozat, hogy teljesülnek az alábbiak:

$$(1) \quad P(L(\mathfrak{g})\text{-nak } \mathfrak{g}_n\text{-ben maximuma van}) \rightarrow 1;$$

$$(2) \quad P(\sqrt{n}\|\mathfrak{g}_n - \mathfrak{g}_0\| < x) \rightarrow P(\beta < x)$$

ahol $\beta \sim N(0, D^{-1})$ eloszlású valószínűségi változó;

$$(3) \quad P(2[L(\mathfrak{g}_n) - L(\mathfrak{g}_0)] < x) \rightarrow P(\chi_x^2 < x).$$

Bizonyítás. 1. Legyen most $\hat{\mathfrak{g}}_n$ a 3.1. tétel szerinti maximumhely és legyen

$$\mathfrak{g}_n = \begin{cases} \hat{\mathfrak{g}}_n; & \text{ha létezik } \hat{\mathfrak{g}}_n \\ 0, & \text{különben.} \end{cases}$$

2. Szintén a 3.1. tétel alapján tudjuk, hogy

$$\|\sqrt{n}(\hat{\mathfrak{g}}_n - \mathfrak{g}_0) - \beta_n\| < \alpha$$

nagy valószínűséggel minden $\alpha > 0$ tetszőleges számra, ha $n > n_0$. De a β_n sorozat aszimptotikusan $N(0, D^{-1})$ eloszlású valószínűségi változó, és mivel $\sqrt{n}(\mathfrak{g}_n - \mathfrak{g}_0)$ β_n -től való eltérése tetszőlegesen kicsiny lehet, ha n elég nagy, így $\sqrt{n}\|\mathfrak{g}_n - \mathfrak{g}_0\|$ is aszimptotikusan $N(0, D^{-1})$ eloszlású valószínűségi változó.

3. A 3.1. tétel szerint egyrészt

$$L(\hat{\mathfrak{g}}_n) = \tilde{L}(\hat{\mathfrak{g}}_n) + \sqrt{n}(\hat{\mathfrak{g}}_n - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) (\hat{\mathfrak{g}}_n - \mathfrak{g}_0) \sqrt{n} + R(\mathfrak{g}),$$

másrészt pedig

$$\lim_{n \rightarrow \infty} \left| \sqrt{n}(\hat{\mathfrak{g}}_n - \mathfrak{g}_0)^T \left(\frac{1}{n} L''(\mathfrak{g}_0) - L \right) (\hat{\mathfrak{g}}_n - \mathfrak{g}_0) \sqrt{n} + R(\mathfrak{g}) \right| \rightarrow 0$$

teljesül.

Láttuk, hogy $\tilde{L}(\hat{\mathfrak{g}}_n) = L(\mathfrak{g}_0) - \frac{1}{2} \beta_n^T L \beta_n$, tehát

$$2[\tilde{L}(\hat{\mathfrak{g}}_n) - L(\mathfrak{g}_0)] = -\frac{1}{2} \beta_n^T L \beta_n = \beta_n^T D \beta_n.$$

D^{-1} pozitív definit szimmetrikus, így van olyan C szimmetrikus mátrix, melyre $C^2 = D$. Ekkor $C^{-1}C^{-1} = D^{-1}$ teljesül. A 3.1. tétel jelöléseivel $\beta_n \rightarrow \beta$ így,

$$\eta_n = C\beta_n \rightarrow \eta = C\beta.$$

Nyilván η is normális eloszlású, $M(\eta) = 0$, továbbá η kovarianciája $CD^{-1}C = CC^{-1}CC^{-1} = I$.

A fentiekből világos, hogy η komponensei függetlenek és standard normálisok. Másrészt $\beta_n^T D \beta_n \rightarrow \beta^T D \beta$ és

$$\beta^T D \beta = \eta^T C^{-1} D C^{-1} \eta = \eta^T C^{-1} C C C^{-1} \eta = \eta^T \eta,$$

amiből

$$\beta_n^T D \beta_n \rightarrow \chi_r^2.$$

IRODALOM

- [1] AITCHISON, J. T. and SILVEY, S. D., "Maximum likelihood estimation of parameters subject to restraints", *Annals of Mathematical Statistics* **29** (1958) 813–828.
- [2] BAKALARCZYK, M., "Asymptotic behaviour of the maximum likelihood estimators for parametric models with contamination", Preprint Nr. 223 Polish Academy of Sciences, 1980.
- [3] BERK, R. H., "Consistency and asymptotic normality of MLE's for exponential models", *Annals of Mathematical Statistics* **43** (1972) 193–204.
- [4] BILINGSLEY, P., *Statistical Inference for Markov Processes* (The University of Chicago Press, Chicago, 1961).
- [5] CHANDA, K. C., "A note on the consistency and maxima of the roots of the likelihood equations", *Biometrika* **41** (1954) 56–61.
- [6] CRAMER, H., *Mathematical Methods of Statistics* (Princeton University Press, 1946).
- [7] CRAMER, H. and WOLD, H., "Some theorems on distribution functions", *Journal of London Mathematical Society* **11** (1936) 290–295.
- [8] DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B., "Maximum likelihood from incomplete data via the EM algorithm", *Journal of Royal Statistical Society, B*, **39** (1977) 1–22.
- [9] DON, F. J. H., "A note on Sheppard's corrections for grouping and maximum likelihood estimation", *Journal of Multivariate Analysis* **11** (1981) 452–458.
- [10] FISZ, M., *Probability Theory and Mathematical Statistics* (Wiley, New York, 1963).
- [11] HSIEK, H. K., "On asymptotic optimality of likelihood ratio tests for multivariate normal distributions", *The Annals of Statistics* **7** (1979) 592–598.
- [12] HUZURBAZAR, V. S., "The likelihood equation consistency and the maxima of the likelihood function", *Annals of Eugenics* **14** (1949) 185–200.
- [13] JUPP, P. E. and MARDIA, K. V., "Maximum likelihood estimators for the matrix of Mises—Fisher and Bingham distributions", *The Annals of Statistics* **7** (1979) 599–606.
- [14] KENT, J. T., "The Fisher—Bingham distribution on the sphere" Research Reports Nr. 7, University of Leeds, 1981.
- [15] KIEFER, J. and WOLFOWITZ, J., "Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters", *Annals of Mathematical Statistics* **27** (1956) 887–906.
- [16] LE CANN, L., "On some asymptotic properties of maximum likelihood estimates and related Bayes estimates", *University of California Publications in Statistics* **1** (1952) 227–330.
- [17] LINDLEY, D. V., "Grouping corrections and maximum likelihood equations", *Proceedings of Cambridge Philosophical Society* **46** (1950) 106–110.
- [18] PETERS, B. C. and WALKER, H. F., "An iterative procedure for obtaining maximum likelihood estimates of the parameters for a mixture of normal distributions", *SIAM Journal on Applied Mathematics* **35** (1978) 362–378.
- [19] PICK, R. and TUSNÁDY, G., "Decomposition of mixtures", *Studia Scientiarum Mathematicae Hungaricae* (1982).
- [20] RAO, C. R., "Maximum likelihood estimation for the multinomial distribution", *Sankhya* **18** (1957) 139–148.

- [21] REDNER, R., "Note on consistency of the maximum likelihood estimate for nonidentifiable distributions", *The Annals of Statistics* 9 (1981) 225—228.
- [22] RÉNYI, A., *Valószínűségszámítás* (Tankönyvkiadó, Budapest, 1973).
- [23] SE, K. P., "Asymptotic properties of maximum likelihood estimators based on conditional specification", *The Annals of Statistics* 7 (1979) 1019—1033.
- [24] WALD, A., "A note on the consistency of the maximum likelihood estimate", *Annals of Mathematical Statistics* 20 (1949) 595—601.
- [25] WOLFOWITZ, J., "On Wald's proof of the consistency of the maximum likelihood estimate", *Annals of Mathematical Statistics* 20 (1949) 601—602.

(Beérkezett: 1982. október 5.)

PICK RÓBERT

1022 BUDAPEST, BIMBÓ ÚT 63. II. 5.

ON THE ASYMPTOTIC OPTIMALITY OF THE MAXIMUM LIKELIHOOD ESTIMATION

R. PICK

In the present paper we investigate the asymptotic properties of the maximum likelihood estimation. By the aid of weakening the Cramer conditions we prove the asymptotic consistency and normality of the method in case of not commonly distributed random variables. So we get a valuable result for investigation of the estimation from practical point of view, too.

HÍREK ÉS KÖZLEMÉNYEK

GAMM-Tagung 1982

Budapest, 1982. április 13—16

A *Gesellschaft für Angewandte Mathematik und Mechanik* (GAMM — *Alkalmazott Matematikai és Mechanikai Társaság*) 1982. évi konferenciáját Budapesten tartotta április 13. és 16. között. A konferencia szervezését a GAMM és a *Bolyai János Matematikai Társulat* végezte a *Magyar Tudományos Akadémia* támogatásával.

A GAMM elsősorban német nyelvterületen működő társaság. Közel 2000 főnyi tagsága között azonban számos ország, így hazánk is képviselteti magát. A GAMM-ot 1922-ben alapították, létrejöttében közvetlenül LUDWIG PRANDTL (*Göttingen*), RICHARD V. MISES (*Berlin*) és HANS J. REISSNER (*Charlottenburg*) tevékenykedett. A társaság megalapításának célját az 1921-ben alapított *Zeitschrift für Angewandte Mathematik und Mechanik* (ZAMM)-ban olvashatjuk: „...folyó év szeptember 21-én a mérnöki tudományok, a mechanika és az alkalmazott matematika képviselői egy társaságot alapítottak, melynek feladata a matematika és a fizika összes, a mérnöki tudományokhoz számítható részterületén folyó kutatás művelése és támogatása.” (ZAMM, 2. 4. füzet, 404. old.). A GAMM alapításának évétől rendszeresen (általában évente) tart konferenciát részben saját, részben társrendezvényként. A budapesti GAMM-Tagung az 54. ilyen konferencia volt. A szocialista országok közül másodikként léptünk be GAMM konferencia rendezői közé (1968-ban *Prágában* volt — először nem német nyelvterületen — GAMM-Tagung). A konferenciákon, melyeken a GAMM tagjai kedvezményes, részvételi díjjal vehetnek részt, tagsági összejevetelt tartanak. A tagsági összejeveteleken választják újá 3 évenként a vezetőséget. A jelenlegi vezetőség 5 tagból áll:

Elnök: Prof. DR. K. KIRCHGÄSSNER (*Stuttgart*).

Alelnök: Prof. DR. P. HENRICI (*Zürich*).

Pénztáros: Prof. DR.-Ing. J. ZIEREP (*Karlsruhe*).

Titkár: Prof. DR. W. VELTE (*Würzburg*).

Titkárhelyettes: Prof. DR. B. BROSKOWSKI (*Frankfurt*).

A konferencia szervezését a GAMM-vezetőség mellett egy szervező bizottság irányította.

Elnök: DR. PRÉKOPA ANDRÁS egyet. tanár, az MTA levelező tagja.

Tagok: DR. BOSZNYAI ÁDÁM egyet. tan., DR. FARKAS MIKLÓS egyet. tan., DR. KALISZKY SÁNDOR egyet. tan., DR. MIKOLÁS MIKLÓS egyet. tan., DR. SZABÓ JÁNOS egyet. tan., az MTA tagja.

Titkárok: DR. BERNAU HEINZ, DR. GÁSPÁR ZSOLT, DR. KOTSIS DOMOKOSNÉ, DR. POPPER GYÖRGY.

A konferenciára 539 résztvevő érkezett 141 kísérő személylyel. Ebből 159 volt a magyar szakember. Összesen 21 ország képviseltette magát, a külföldiek közül kimagaslóan legnagyobb létszámmal a *Német Szövetségi Köztársaság* (284 résztvevővel) (a GAMM tagok túlnyomó többsége is innen való), de a 17 osztrák, 17 csehszlovák, 10 lengyel, 8 jugoszláv résztvevő mellett többen érkeztek a *Szovjetunióból*, az *NDK-ból*, *Nagy-Britanniából* is, s volt résztvevő az *Amerikai Egyesült Államokból*, *Indiából* és *Venezuelából* is.

A konferencia előadásait a *Magyar Tudományos Akadémia* központi épületében, valamint az *MTESZ Kossuth Lajos téri székházában* tartották. A konferenciát DR. PRÉKOPA ANDRÁS, az MTA lev. tagja, a szervező bizottság elnöke, a *Bolyai János Matematikai Társulat Alkalmazott Matematikai Szakosztályának elnöke* nyitotta meg. Bevezető szavaiban értékelte annak jelentőségét, hogy a GAMM Budapesten rendezte ez évi konferenciáját. Megemlékezett a magyar alkalmazott matematikusok és mechanikusok, valamint a GAMM közismerten jó kapcsolatáról, majd utalt a magyar

tudósok (NEUMANN, RÉNYI, JORDAN, EGERVÁRY, valamint KÁRMÁN és REUSS) ezen területen elért eredményeire, akik közül egyesek éppen az alkalmazott matematika és a mechanika egységes szemléletmódját tudták megteremteni, mely a GAMM-nak is fontos célkitűzése. Külön kiemelte annak jelentőségét, hogy a GAMM-ban olyan modern tudományágak, mint az operációkutatás és a számítógéptudomány, egyre inkább tért nyernek.

Másodikként DR. GESZTI P. OTTÓ, egyetemi tanár, a *Magyar Tudományos Akadémia Műszaki Tudományok Osztályának elnöke* üdvözölte a konferenciát az MTA nevében. A GAMM részéről Prof. DR. K. KIRCHGÄSSNER, a GAMM elnöke válaszolt az üdvözlésekre. Beszédében méltatta a magyar tudósok szerepét az egyetemes kultúrtörténetben, részletesen megemlékezett BOLYAI JÁNOS munkásságáról, majd megköszönve a szíves fogadtatást, ő is sikeres munkát kívánt a konferencia valamennyi résztvevőjének.

A konferencia délelőtti részén az *Akadémia Dísztermében* hangzottak el a — részben egy-egy tudományterületet áttekintő, részben egy-egy új kutatási eredményt ismertető — plenáris előadások. A plenáris előadások az elhangzás sorrendjében a következők voltak:

- L. COLLATZ (*Hamburg*): *Numerische Anwendungen der Monotonie bei Differentialgleichungen.*
 O. M. BELOCEKOVSKI (*Moszkva*): Прямое статистическое моделирование задач аэродинамики.
 W. SCHIEHLEN (*Stuttgart*): *Mechanische Zufallsschwingungen.*
 J. ROSENMÜLLER (*Bielefeld*): *Cooperative games and nondegeneracy-problems.*
 TH. LEHMANN (*Bochum*): *Einige Aspekte der Thermoplastizität.*
 D. GROSS (*Darmstadt*): *Konzepte der Bruchmechanik.*
 A. TONDL (*Praha*): *Über die Anwendungsmöglichkeiten eines Schwingungsdämpfers in selbst-erregten Rotorsystemen.*
 R. KLÖTZLER (*Leipzig*): *Global optimization in control theory.*
 S. KALISZKY (*Budapest*): *Dynamical loaded plastic structures.*
 R. JELTSCH (*Aachen*): *Stability, contractivity, and accuracy of numerical methods for stiff initial value problems.*
 R. SCHASSBERGER (*Berlin*): *Networks of stochastic servers.*

Közbejött akadályok miatt nem tudott eljönni, s így nem tudta megtartani T. GARZA (*Mexico*) "On statistical methods in seismology" c. előadását. Nem tudott részt venni a konferencián két további meghívott plenáris előadó sem, W. FISZDON *Lengyelországból* és Sz. M. NYIKOLSZKI a *Szovjetunióból* (tervezett előadásuk címét előre nem közölték).

A GAMM alapítójának tiszteletére a GAMM és a *Deutsche Gesellschaft für Luft- und Raumfahrt* (DGLR) közös rendezésében Prof. DR. W. WUEST (*Institut für Experimentelle Strömungsmechanik der DFVLR Göttingen*) tartotta meg 13-án délelőtt a hagyományos, sorrendben a 25. LUDWIG PRANDTL emlékelőadást "Thermokapillare Strömungen (Thermische Marangonikonvektion)" címmel. 15-én 19 órakor rendeztek meg a GAMM hagyományához híven egy ún. népszerű tudományos előadást. A nagyszámú hallgatóság DR. FARKAS MIKLÓS egyet. tan. (BME) "Mathematics and objective reality" című előadását figyelte nagy érdeklődéssel. Az előadó a matematika helyét és szerepét taglalta, különös tekintettel a többi tudományokhoz való viszonyára.

A délutáni szekcielőadásokat a konferencia két színhelyén, összesen 7 szekcióban tartották meg. A GAMM az utóbbi évek konferenciáin a következő témakörökben fogad el rövid előadásokat: 1. Merev testek mechanikája, 2. Lengés- és stabilitáselmélet, 3. Rugalmasságtan és képlékenységtan, 4. Folyadékok mechanikája, 5. Alkalmazott analízis és matematikai fizika, 6. Numerikus analízis, 7. Optimalizálás, sztochasztika és matematikai módszerek a közgazdaságtudományban, 8. Számítógéptudomány. (A 8. szekció az utóbbi években többször hiányzott, pl. az 1981-es *würzburgi* GAMM-Tagung is csak 7 szekciót hirdetett meg.) Az 1. szekcióba nem érkezett előadás. A 2—4. szekció előadásait az MTESZ székházban, az 5—8. szekció előadásait az MTA épületében bonyolították le. Összesen 347 rövid előadás hangzott el. A 8. szekció ismételt felvétele nem látszott túlzottan előkészítettnek, ebbe a szekcióba csupán 4 előadás érkezett. A többi szekcióban 50—70 előadás hangzott el. Az előadások száma, valamint színvonala minden tekintetben megfelelt az eddigi GAMM konferenciáknak, az ún. hagyományos szekciók népszerűsége mellett igen nagyszámú és színvonalas előadás hangzott el az Optimalizálás szekcióban. A szekcióülések elnöki teendőit neves külföldi és hazai tudósok látták el. A hozzászólások nagy száma az érdeklődés mértékét jellemezte. Külön megbeszélést tartottak az érdeklődők a végeelem módszerrel kapcsolatban.

A tudományos programmal részben párhuzamosan folytak a kulturális rendezvények. Első este fogadás volt valamennyi résztvevő és kísérő részére a *Budapest Szálló éttermében*, 14-én pedig

a *Mátyás templomban* volt orgonakoncert, mely elsősorban a hely vonzereje miatt jelentett élményt. Az IBUSZ által rendezett programok között válogathattak a résztvevők, szerveztek sétát a *Budai Várba*, kirándulást *Szentendrére* és *Hortobágyra*, valamint a búcsú estén borkostolót. Az MTA a GAMM vezetőség tiszteletére rendezett fogadást.

Hogy a konferencia valóban jól sikerült, azt nemcsak a helyszínen elhangzott pozitív vélemények, hanem neves professzorok — a szervezőkhöz azóta írt — elismerésről tanúskodó levelei is bizonyítják.

KOTSIS DOMOKOSNÉ

A kiadásért felel az Akadémiai Kiadó igazgatója
Műszaki szerkesztő: Sándor István
A kézirat nyomdába érkezett: 1983. február 4. — Terjedelem: 19,25 (A/5 ív)
83-677 — Szegedi Nyomda — Felelős vezető: Dobó József igazgató

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámmal kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék mindig az utolsó szakasz kell hogy legyen, és azt nem kell sorszámmal ellátni. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos postai címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé irt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segéd tételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német, francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozat ábráit és az esetleges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától független, folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként feltüntetett, ábraazonosító sorszámmal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve társszerzők esetén az első szerző neve szerinti alfabetikus sorrendben úgy, hogy külön, de folytatólagos sorszámozású listát alkossanak a latin és a cirill betűs nevű szerzők műveire vonatkozó hivatkozások, és mindkét részben a megfelelő alfabetikus sorrend legyen kialakítva. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., »Über die Theorie der einfachen Ungleichungen«, *Journal für die reine und angewandte Mathematik* 124 (1902) 1—27.
- [2] Kéri, G., „DUALSIMP“, rutin a CDC 3300-as gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertető 2. 1973. május) 19—20.
- [3] Prékopa, A., „Sztohasztikus rendszerek optimalizálási problémáiról“, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U., „Recent research on the ruin problem of collective risk theory“, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam—London, 1973) 221—228.
- [5] Zoutendijk, G., *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76—78]. A szerzők a dolgozatukról 100 darab különlenyomatot kapnak, ezek költsége — nyomott oldalanként 25 forint — a szerzői díjat terheli.

TARTALOMJEGYZÉK

<i>Farkas Miklós: Matematika és objektív valóság</i>	1
<i>Demetrovics János, Füredi Zoltán és Katona Gyula: A függőségek és az individuumok száma közötti kapcsolat összetett adatrendszerek esetén</i>	13
<i>Füredi Zoltán: Minimális relációs adatbázisok</i>	23
<i>Galántai Aurél: Egylépéses módszercsaládok konvergencia- és hibaanalízise</i>	29
<i>Sparing László: A perturbált van der Pol-differenciálegyenlet számítógépes vizsgálata</i>	43
<i>Fülöp János: Eredeti csúcspont keresése és alkalmazása konkáv függvény lineáris feltételek melletti minimalizálásakor</i>	51
<i>Rapcsák Tamás: A nemlineáris programozási feladat optimalitási feltételeinek differenciál-geometriai vizsgálata</i>	73
<i>Komáromi Éva: Logaritmikusan konkáv eloszlásfüggvénnyel korlátozott lineáris programozási feladat megoldásának duális megközelítése</i>	85
<i>Pintér János: A Csebisev-egyenlőtlenség egy élesítése Monte Carlo-módszerrel becsült függvényértékek esetén</i>	93
<i>Bíró Miklós: A legerősebb próbák témakörének matematikai programozási megközelítése</i>	105
<i>Bíró Miklós: A bináris hátizsák feladat</i>	113
<i>Gyires Béla: Egy mátrixegyenlet megoldásáról és az eredmények alkalmazásáról valószínűségi eloszlásfüggvények lineáris kombinációinak elméletében</i>	137
<i>Vetier András: Szinaj-biliárd potenciáltérben, horociklus-konstrukció</i>	145
<i>Tóth János, Érdi Péter és Török László Turul: A Poisson-eloszlás jelentősége összetett kémiai reakciók sztochasztikus modelljében</i>	175
<i>Huhn Edit: Kálmán-szűrők aszimptotikus viselkedéséről</i>	197
<i>Pick Róbert: A maximum likelihood becslés aszimptotikus optimalitásáról</i>	203
<i>Hírek és közlemények</i>	217

INDEX

<i>Farkas, M., Mathematics and objectiv reality</i>	1
<i>Demetrovics, J., Füredi, Z. and Katona, G. O. H., The relation between the number of individuals and dependencies in the composed data bases</i>	13
<i>Füredi, Z., Minimal relational data bases</i>	23
<i>Galántai, A., Convergence theorems and error analysis for the families of one-step methods</i> ...	29
<i>Sparing, L., Numerical investigation of the perturbed van der Pol's equation</i>	43
<i>Fülöp, J., Finding original vertex and its application minimizing a concave function subject to linear constraints</i>	51
<i>Rapcsák, T., The optimality conditions of the nonlinear programming problems and differential geometry</i>	73
<i>Komáromi, É., A dual approach to stochastic linear programming problems constrained by logarithmic concave joint probability distribution function</i>	85
<i>Pintér, J., An improved Chebyshev-inequality for function value estimates by Monte Carlo techniques</i>	93
<i>Bíró, M., On a mathematical programming approach to the topic of the most efficient tests</i> 105	
<i>Bíró, M., The binary knapsack problem</i>	113
<i>Gyires, B., On the solution of certain matrix equation and an application of the results in the theory of linear combinations of probability distribution functions</i>	137
<i>Vetier, A., Sinai-billiard in potencial field</i>	145
<i>Tóth, J., Érdi, P. and Török, T. L., Significance of the Poisson distribution in the stochastic model of complex chemical reactions</i>	175
<i>Huhn, E., On the asymptotic behaviour of Kalman filters</i>	197
<i>Pick, R., On the asymptotic optimality of the Maximum likelihood estimation</i>	203
<i>Communications</i>	217

Alkalmazott matematikai lapok

1983/3-4

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

9.

KÖTET

AKADÉMIAI KIADÓ, BUDAPEST

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI
TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

FŐSZERKESZTŐ

PRÉKOPA ANDRÁS

FŐSZERKESZTŐ-HELYETTES

ARATÓ MÁTYÁS

A SZERKESZTŐ BIZOTTSÁG TAGJAI

BENCZUR ANDRÁS, CSISZÁR IMRE, FARKAS MIKLÓS, GYIRES BÉLA,
HATVANI LÁSZLÓ, HEPPES ALADÁR, KÁTAI IMRE, KIS OTTÓ,
SARKADI KÁROLY, TANDORI KÁROLY, VARGA LÁSZLÓ,
SZÁNTAI TAMÁS (technikai szerkesztő)

MUNKATÁRSAK

BAJCSAY PÁL, BALLA KATALIN, BÉKÉSSY ANDRÁS, CSÁKI PÉTER,
CSIRIK JÁNOS, DEMETROVICS JÁNOS, DÉNES JÓZSEF, DÓMÖLKI BÁLINT,
ELBERT ÁRPÁD, FORGÓ FERENC, GÉCSEG FERENC, GERGELY JÓZSEF,
GESZTELYI ERNŐ, GYÖRFFY LÁSZLÓ, KLAFSZKY EMIL, KÓSA ANDRÁS,
KOVÁCS LÁSZLÓ BÉLA, LÁSZLÓ ZOLTÁN, MIKOLÁS MIKLÓS,
MOGYORÓDI JÓZSEF, NÉMETH GÉZA, NEMETZ TIBOR, RÉVÉSZ PÁL,
RÓZSA PÁL, STAHL JÁNOS, SZÉP JENŐ, TANKÓ JÓZSEF, TOMKÓ JÓZSEF,
TÓKE PÁL, TUSNÁDY GÁBOR, VINCZE ENDRE

IX. kötet 3—4. szám

Szerkesztőség: 1502 Budapest XI., Kende u. 13—17.

Kiadóhivatal: 1055 Budapest V., Alkotmány u. 21.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredmények számítógépes programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztő bizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Prékopa András, főszerkesztő
1502 Budapest, Kende u. 13—17.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 100 forint. Belföldi megrendelések az Akadémiai Kiadó, 1055 Budapest V., Alkotmány u. 21. címen (pénzforgalmi jelzőszám 215—11 488), külföldi megrendelések a Kultúra Külkereskedelmi Vállalat, H-1389 Budapest, Pf. 149. címen (pénzforgalmi jelzőszám 218—10 990) lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungaricae,
2. Acta Physica Hungaricae,
3. Studia Scientiarum Mathematicarum Hungarica.

NAGYMÉRETŰ, VEGYESVÁLTOZÓS, MATEMATIKAI MODELL TERMÍKUS VILLAMOSENERGIA-RENDSZER RÖVIDTÁVÚ, OPTIMÁLIS MENETRENDJÉNEK MEGHATÁROZÁSÁRA HÁLÓZATI FELTÉTELEK FIGYELEMBEVÉTELÉVEL¹

DEÁK ISTVÁN, HOFFER JÁNOS, MAYER JÁNOS, NÉMETH ÁGOSTON,
POTECZ BÉLA, PRÉKOPA ANDRÁS, STRAZICKY BEÁTA

Budapest

Ez a szokásosnál nagyobb terjedelmű cikk az MTA SZTAKI Operációkutatási Osztályán végzett, a villamosenergia-termelő rendszer optimális napi ütemezését célzó sokéves munka első részletes be számolója.

A munka történetét ismertető bevezetés és a villamosenergia-rendszerek irányítási feladatait áttekintő 2. fejezet után a cikk 3. fejezete termikus erőművekből álló energiatermelő rendszer egy általános matematikai programozási modelljét tartalmazza. A modell általános abban az értelemben, hogy megfogalmazásánál nem tettünk egyszerűsítő feltevéseket a modellnek megfelelő nemlineáris, vegyes (folytonos-diszkrét) változójú feladat megoldhatósága érdekében. A modell tartalmazza az erőművi viszonyokat leíró feltételek mellett az elektromos hálózat zavartalan működéséhez szükséges feltételeket is. Minimalizálandó célfüggvénye az energiatermelés költsége, mely a tüzelőanyagfel- használás, az erőművi egységek állás-újraindítás költségéből és a hálózati veszteség költség összegé- ből tevődik össze. A 4. fejezetben az ütemezési problémát egyszerűsítő feltevéseket ismertetjük, majd az 5. fejezet az említett feltevések figyelembevételével kapott modell leírása. Ennek az egy- szerűsített modellnek nagyméretű, speciális struktúrájú, vegyesváltozós, lineáris programozási fela- dat felel meg, melynek megoldásáról a 6. fejezetben számolunk be.

A terjedelmes, három részből álló függelék a villamosenergia-rendszerek átviteli hálózatának alapismereteit és összefüggéseit foglalja össze olyan formában, mely — megítélésünk szerint — ma- tematikusok számára könnyen érthető, és a matematikai számításokhoz jó alapul szolgál.

1. Bevezetés

Ez a dolgozat egy kutató csoport sokéves munkájának első részletes beszámolója. A munka az MTA SZTAKI Operációkutatási Osztályán készült.

Több mint 10 évvel ezelőtt, 1972-ben, a szerzők egyike — PRÉKOPA ANDRÁS — Olaszországban járt, előadásokat tartott a Római Egyetemen. Az előadásokon mér- nökök is részt vettek. Egyikük, abban a reményben, hogy az előadott matematikai programozási modellek számos villamosenergia-ipari alkalmazást is nyerhetnek (miután vízügyi alkalmazásokról sok szó esett), elkauzolta az előadót a római teher- elosztóba. Ennek egyik kutatója korábban már kidolgozott egy dinamikus típusú modellt a villamosenergia-termelés napi ütemezésének a meghatározására. A kevés változós, nagy aggregáltsági fokú modell csupán globális értékeket adott, ám abban az időben más teherelosztók sem rendelkeztek még finomabb időbeli és térbeli fel- osztást tartalmazó, napi menetrend meghatározását nyújtó matematikai programo- zási modellel. Ekkor született meg a gondolat, hogy meg kellene próbálni a napi me- netrend meghatározására megalkotni egy nagyméretű, vegyesváltozós modellt a hazai körülményeknek megfelelően, de azért az általánosság némi igényével.

¹ Készült az MTA Központi Kutatási alapjára kötött szerződés keretében.

A gondolatot követő első tett abban állt, hogy a téma iránt sikerült felkelteni NÉMETH ÁGOSTON erősáramú villamosmérnök hallgató érdeklődését, aki azután erről írta diplomamunkáját, melyet 1973-ban védett meg. Ez lényegében a szakirodalomban ismert eljárások egy részét foglalta össze. Miután POTE CZ BÉLA, az Országos Teherelosztó akkori osztályvezetője, mai igazgatóhelyettese, a diplomamunkával megismerkedett és tudomást szerzett a további tervekről, megszerezte a Magyar Vilamos Művek Tröszt támogatását és személyesen is bekapcsolódott a munkába.

Az első modell, mely az ennek a dolgozatnak az 5. fejezetében taglalt modell lényeges elemeit már tartalmazta, NÉMETH ÁGOSTON és PRÉKOPA ANDRÁS munkájaként született meg POTE CZ BÉLA hathatós közreműködésével. Ez 1974 végére készült el. Az eredményeket egy nem publikált tanulmány foglalta össze [35]. A fő eredmény ebben annak a vegyesváltozós modellnek az összeállítása, melynek fejlettebb formája e dolgozat 5. fejezetében olvasható.

A [35] tanulmánynak a szerzői között az eddig említett három személyen kívül szerepelt még KÉRI GERZSON, TURCHÁNYI PIROSKA és VÍZVÁRI BÉLA is. KÉRI GERZSON összefoglaló áttekintést nyújtott az itthon elérhető lineáris programozási programcsomagokról, TURCHÁNYI PIROSKA és VÍZVÁRI BÉLA pedig egy-egy kisméretű, folytonos változós, illetve diszkrét változós mintafeladatot oldott meg, meglevő, részben általuk készített általános célú programok segítségével.

A munkának ebben az első fázisában részt vett még MOLNÁR MIKLÓS közgazdász is, aki nagy segítséget nyújtott a megfelelő közgazdasági szemlélet kialakításában. Ami azonban a modellalkotási munkát illeti, ő különállóan megfogalmazott egy modellt, mely ugyan elvi szempontból érdekes volt, azonban a gyakorlati számításokra nem volt megfelelő.

A munka második fázisa 1976-ban kezdődött. Ekkor kapcsolódott be STRAZICKY BEÁTA. Az 1976–77 években fő feladata a [35] tanulmányban foglalt vegyesváltozós modell számszerűsítése volt. Tisztázta, hogy milyen adatok állnak rendelkezésre, ezek milyen formában adóttak és hogyan építhető fel ezekből az egy konkrét napnak megfelelő feladat. Ezekben az években készítette el az ELTE matematika (operációkutatási) szakon HOFFER JÁNOS szakdolgozatát, a *Benders-féle dekompozícióról*, STRAZICKY BEÁTA vezetésével.

A feladat számszerűsítési munkáinak keretén belül 1977–78-ban BÁRÁSZ ANDRÁS közreműködésével elkészült a Magyar Tudományos Akadémia CDC 3300-as számítógépére a feladathoz szükséges adatbázis.

A napi menetrend feladatának (a mostani cikk 5. fejeztében foglalt feladatról van szó) numerikus megoldására vonatkozó algoritmus kialakítása, valamint a kapcsolódó számítógépes programok elkészítése az 1977–1981 években történt. Hoszszas megfontolás után a *Benders-féle dekompozíció* alkalmazása mellett döntöttünk. Először STRAZICKY BEÁTA elkészítette a feszültséges feltételek nélküli feladat generálására vonatkozó programokat. HOFFER JÁNOS *Benders dekompozícióval* megoldotta a három különböző típusú periódus feladatát és elkészítette a megfelelő számítógépes programokat.

MAYER JÁNOS 1978-ban kapcsolódott be a munkába. Ő az ún. feszültséges periódusok villamos feltétel rendszerét generáló és a „load flow feladat” megoldását célzó programrendszer kifejlesztésén dolgozott. 1979 végére elkészítette a hálózati feltételeket generáló módszert és a megfelelő számítógépes program első változatát. A szakirodalom tanulmányozása után kiválasztotta a később nagyon jónak bizo-

nyult *Stott-féle módszert*, melynek implementálását 1981-ben fejezte be. 1980-ban elkészítette a feszültséges feltételeket generáló módszer programjait és a megfelelő adatszervezési programokat. A feszültséges feltételek tesztelésében HOFFER JÁNOS is részt vett. A teljes menetrendet előállító program oly módon született, hogy a feszültséges feltételeket generáló programot HOFFER JÁNOS és MAYER JÁNOS illesztette STRAZICKY BEÁTA fent említett feladat generáló programjához.

DEÁK ISTVÁN is 1978-ban kapcsolódott be a munkába. Ő készítette az optimalizáló program részeinek összefogását biztosító keretprogramot, továbbá a tüzelőanyagkényszer feltételt is figyelembe vevő megoldást előállító programot. Ennek az utóbbinak az elve POTECZ BÉLÁTÓL származik.

A teljes program első változata 1979-ben készült el, és egy 10 csomópontos, kis-méretű tesztfeladaton jól le is futott. A program futtatása azonban teljes kudarcot vallott reális méretű, 170 csomópontos hálózatokon. Ennek, mint a későbbiekben kiderült, kettős oka volt. Egyfelől rendkívül sok és változatos hiba volt az adatokban, másrészt a *load flow feladat* megoldására kialakított saját módszerünk nagyméretű feladatok esetében nem konvergált. A módszer hiányosságainak a felfedezésében nagy segítségünkre volt BORZSÁK PÉTER villamos hálózatokra vonatkozó szakértelme. Ami az adatok hibáinak a felfedezését illeti, ez MAYER JÁNOSTÓL vett el hosszú hónapokat, annak ellenére, hogy NÉMETH ÁGOSTON e téren is komoly segítséget nyújtott. Szinte egyenként kellett a hibákat kiszűrni a modell futtatásaira és az eredmények analizésére támaszkodva. Az adatjavítás során a modell több kisebb módosításon ment át (ilyen módosítás volt pl. az, hogy egy csomópont-hoz több erőmű is csatlakozhat).

Meg kell még említeni, hogy mind a *load flow feladat* megoldása, mind pedig a feszültséges feltételek generálása szükségessé tette egy ritka mátrixú, nagyméretű lineáris egyenletrendszer megoldását. Ehhez 1978-ban KÉRI GERZSON ajánlott egy ritka mátrix invertáló, általa készített, általános célkitűzésű módszert és programot, melyet azonban később mégsem építettünk be, minthogy ez a program az inverz mátrixot állítja elő, és ezért egyenletrendszer-megoldási célra lassúnak bizonyult. A program megbízhatóan működött, a mi speciális feladatunkra való gyakori alkalmazása azonban nem bizonyult gazdaságosnak.

Az általunk ténylegesen felhasznált, lineáris egyenletrendszert megoldó programot a feladat specialitásainak figyelembevételével SPARING LÁSZLÓ készítette el, ez volt a diplomamunkája a BME matematikus-mérnök szakán. Ezt MAYER JÁNOS és PRÉKOPA ANDRÁS vezetésével készítette.

SPARING LÁSZLÓ munkája lehetővé tette, hogy MAYER JÁNOS implementálja *Stott módszerét* a *load flow feladat* megoldására. Az eredmények előírt formában való táblázatosítására vonatkozó programokat FÜLÖP JÁNOS készítette el.

Az optimális napi menetrend meghatározását célzó munka második szakasza 1981-ben fejeződött be. A feladat gépi-numerikus megoldása reális méretben, reális adatokkal ellátott tesztfeladatokon sikeresen lefutott a CDC 3300-as számítógépen. Az addigi eredményekről az ugyancsak nem publikált [11] tanulmány számolt be.

A feladat sikeres megoldása utáni munkálatok az alábbi célkitűzésekkel folytak:

- a programokat átirni az Akadémia IBM 3031-es számítógépére,
- az IBM 3031-es adta lehetőségeken túlmenően is csökkenteni a futási időt, új adatbázist kialakítani,

- a napi menetrend feladatának elvi alapjait, továbbá az általunk adott megoldást egységes matematikai keretbe foglalni.

A dolgozat nem foglalkozik azzal, hogy milyen eredmények születtek az első két cél elérése érdekében tett munkák során, viszont meg kíván felelni az utolsó célkitűzésnek. Az alábbiakban rövid tartalmi összefoglalást adunk.

A 2. fejezet a villamosenergia-rendszerekkel kapcsolatos, a matematikai modellalkotáshoz szükséges legfontosabb műszaki ismereteket foglalja össze. Minthogy ezek együttese a feladat fizikai értelemben vett meghatározását nyújtja, nevezhetjük a 2. fejezetben foglaltakat a feladat fizikai modelljének is.

A 3. fejezetben elvonatkoztatunk a magyar villamosenergia-ipar konkrét feladatától és megfogalmazzuk a villamosenergia-termelés ütemezésének általános feladatát, amelyben csupán termikus erőművek vesznek részt az energia termelésében. Az erőművek üzemmódjaira épülő vegyesváltozós modellben fontosnak tartjuk azt, hogy a hálózati feltételeket is szerepeltetjük a modellben.

A 4. fejezetben szisztematikusan összeállítottuk azokat az egyszerűsítő feltevéseket, amelyek szerepet játszanak az előző fejezet modelljének a magyar villamosenergia-rendszerre való alkalmazásában. Ezek a feltevések a nyert feladat gépi-numerikus megoldásának lehetővé tételét célozzák.

Az 5. fejezetben megadjuk a vegyesváltozós, egyszerűsített modellt, melynek célfüggvénye és feltételei lineárisak, együttható mátrixa pedig strukturált.

A 6. fejezet a feladat megoldásával foglalkozik. A módszer a *Benders-dekompozíció*n alapul, heurisztikus elemeket is tartalmaz, és a feladat szerkezetét lényegesen kihasználja.

A terjedelmes, négy részből álló függelék a villamosenergia-rendszerek átviteli hálózatának alapismereteit és összefüggéseit foglalja össze olyan formában, mely — megítélésünk szerint — matematikusok számára könnyen érthető, és a matematikai számításokhoz jó alapul szolgál.

Nem tartalmazza a dolgozat a feladat megoldásához készített programrendszer ismertetését, és futási eredményekről szóló részletes beszámolókat sem mellékelünk. Ezek a [11] belső intézeti anyagban találhatók meg, ám miután a programok hamarosan futnak majd az IBM 3031-es gépen is, ezek a korábbi programleírások és futási eredményekről szóló beszámolók aktualitásukat veszítik.

A fenti leírásból kitűnik, hogy az idők folyamán igen sokan részt vettek az optimális napi menetrend kialakítását célzó projekt munkálataiban. Mégis, a jelenlegi dolgozat szerzői azok, akik a legfontosabb munkákat elvégezték. Mellettük SPARING LÁSZLÓ játszott még jelentős szerepet, az ő hozzájárulása azonban jól körülhatárolható részfeladat megoldása volt, és eredményeit részletesebb formában külön fogja publikálni.

Nehéz dolog az egyes fejezetek szerzőinek a nevét egyértelműen megadni, hiszen a modell sok közös megbeszélésen alakult, formálódott, azonkívül ennek a dolgozatnak a fő modellje már az első, 1974-es [35] beszámolóban is megtalálható szerényebb formában, és más személy volt a szerző a későbbiek számára alapul szolgáló műben, mint aki ennek a dolgozatnak a megfelelő fejezetét írta. Ilyenformán az alábbi fel-

sorolásban sokkal inkább a fejezetek szerzőit soroljuk fel, mintsem azokat, akik az egyes fejezeteket jelenlegi végleges formájukban megírták.

A 2. fejezet főként NÉMETH ÁGOSTON és kisebb részben PRÉKOPA ANDRÁS munkája.

A dolgozat architektúrájának a kialakításában meghatározó szerepe volt PRÉKOPA ANDRÁSNAK, aki kezdeményezte az 5. fejezetben foglalt speciális modell mellett a 3. fejezetben foglalt általános modell kialakítását, valamint a villamosenergia-rendszerek átviteli hálózatának a Függelékben foglalt tárgyalásmódját is. A korábbi modell megfogalmazása, valamint e munkája révén társszerzője a dolgozat valamennyi fejezetének a 6. kivételével.

A villamosenergia-rendszerekkel kapcsolatos ismeretek közvetítése és a modellalkotás révén NÉMETH ÁGOSTON is társszerzőnek tekinthető a 3., 4. és az 5. fejezeteket illetően. Ugyanez vonatkozik POTE CZ BÉLÁRA is, akinek a gondolatai a 6. fejezetben is lényeges szerephez jutottak.

A 3., 4. és 5. fejezetek további szerzői és a jelenlegi végleges változatuk kialakítói STRAZICKY BEÁTA és — a feszültséges feltételek vonatkozásában — MAYER JÁNOS.

Mint korábban említettük, a vegyesváltozós feladat megoldásában POTE CZ BÉLA gondolatai lényeges szerepet játszottak. A 6. fejezet anyaga ezek felhasználásával készült, azonban főként DEÁK ISTVÁN és HOFFER JÁNOS munkája eredményeként.

Végül a Függelék nagyrészt MAYER JÁNOS munkájaként könyvelhető el.

A szerzőknek a fejezetekhez történt fenti hozzárendelése — ismételten hangsúlyozzuk — csak hozzávetőleges, mert az egyes szerzők munkái nem választhatók szét egyértelműen.

Most néhány gondolatot fejtünk ki arra vonatkozólag, hogy modellünk hogyan illeszthető a matematikai programozási modellek rendszerébe, milyen motivációk merültek fel a matematikai módszertan oldaláról.

Időben lejátszódó folyamatok modellálása és irányítása speciális struktúrájú, nagyméretű matematikai programozási feladatok segítségével, régi keletű elképzelés. DANTZIG a szimplex módszert 1947-ben alkotta meg, és 1951-ben publikálta. Ám, már 1949-ben két cikket írt az ún. lépcsős struktúrájú, nagyméretű lineáris programozási feladatokról, melyek az *Econometrica* hasábjain jelentek meg. Az egyiket M. K. WOODDAL közösen írta, és ebben a berlini léghíd megszervezése szerepel illusztratív alkalmazásként. A második dolgozatban a *dinamikus típusú Leontief-féle input-output modellel* való kapcsolatot fejtegeti a szerző.

Ha nem is lépcsős struktúrájú, de speciális szerkezetű (ún. blokk-trianguláris) mátrix szerepel NEUMANN JÁNOSNAK a konstans módon fejlődő gazdaságra vonatkozó matematikai modelljében.

A speciális struktúrájú lineáris, vagy nemlineáris programozási feladatok nagy szerephez jutnak a sztochasztikus programozásban. Ilyen esetekben egy kiindulásul szolgáló feladat mátrixa sokszorozódik meg egy nagyméretű mátrixon belül kb. annyiszor, mint amennyi lehetséges értéke van a feladatban szereplő véletlen mennyiségeknek együtt. Az ezzel kapcsolatos első eredmények szintén DANTZIGTÓL származnak, és a több mint 20 évvel ezelőtti időben keletkeztek.

A nagyméretű matematikai programozási feladatok gépi-numerikus megoldhatósága bátorítólag hat ilyen jellegű modellek alkotására. Minthogy a változók és a feltételek száma igen nagy lehet (ezres nagyságrendű feladat még reális idő alatt megold-

ható), ezért az időbeli folyamatok nem elnagyolt, hanem kellő részletességű, realiztikus modellálására és irányítására lehetőség nyílt.

Ez adta a gondolatot és a bátorságot az ebben a cikkben leírt modell megfogalmazására.

A munka első fázisának befejezésekor azonban kitűnt, hogy modellünk jelentősen eltér a lineáristól nemcsak azáltal, hogy a változók egy (nem csekély) része diszkrét, 0—1 értékű, hanem azáltal is — és a fő nehézséget tulajdonképpen ez okozta —, hogy az erőművekben lejátszódó folyamatok modellálásán kívül a villamos hálózat adottságait is figyelembe kell vennünk, ez pedig azt jelentette, hogy a feltételek jelentős része nemlineáris, sőt ezek explicite nem is adottak a rendelkezésre álló adatok alapján, hanem bonyolult ágterhelési feladatok megoldása révén kell azokat generálnunk.

Ezek után egyelőre félre kellett tenni a felmerülő sztochasztikus programozási szempontokat és a determinisztikus feladatra kellett koncentrálnunk. A sztochasztikus szempontok az egynapos modellben különben sem jutnak érvényre, hiszen egyfelől az egy napi villamosenergia-igény kielégítő pontossággal előre jelezhető, és az ebben rejlő véletlen hiba a modellhiba nagyságrendjén belül van. A későbbiekben azonban foglalkozni kívánunk olyan modell konstrukciójával is, mely többnapos időszakra vonatkozik. Ebben a sztochasztikus programozás minden bizonnyal szerepet fog majd kapni.

Az alábbiakban rövid irodalmi áttekintést adunk.

Cikkünk témája több szakterülethez kapcsolódik, ezért az Olvasó tájékozódását megkönnyítendő, az irodalomjegyzékben néhány könyvet is feltüntetünk az egyes területekre vonatkozóan.

A villamosenergia-rendszerekkel kapcsolatos műszaki témákkal foglalkozik [7], [19], [22], [23], [46] és [73]. Villamos hálózatokat tárgyalnak a [60], [69] könyvek; az átviteli hálózatra vonatkozó hálózatszámítási feladattal és annak megoldásával kapcsolatban a [3], [19], [20], [23], [63], [67] könyveket ajánljuk.

Az „Együttműködő erőművek gazdaságos irányítása” témakörben a [19], [36], [37] művek szolgálhatnak tájékozódásul.

A hazai kutatási eredményekre vonatkozóan lásd a [70] tanulmányt, annak irodalomjegyzékében felsorolt dolgozatokat, valamint a [71], [53], [68] tanulmányokat.

A megoldásra alkalmazott matematikai, operációkutatási módszerekre vonatkozóan a [9], [40], [43], [55], [57] könyveket ajánljuk.

Villamosenergia-rendszerek átviteli hálózatával kapcsolatban a leggyakrabban felmerülő numerikus probléma a hálózatszámítási (*load flow*) feladat megoldása. A megoldó algoritmusokra vonatkozóan az 1973-ig terjedő időszak eredményeinek összefoglalása megtalálható a [66] dolgozatban. Az azóta kidolgozott módszerek közül a gyakorlatban legjobban bevált algoritmusok egyike: *Stott módszere* [65], melyet a dolgozat függelékében részletesen tárgyalunk.

Az optimális teljesítményáramlás problémakör olyan időszakra vonatkozik, melyben a teljesítményigény konstansnak tekinthető. A feladat: meghatározni az átviteli hálózat egyes csomópontjaihoz csatlakozó erőművek által termelendő teljesítményeket oly módon, hogy a tüzelőanyag költség vagy a hálózati veszteség minimális legyen. Korlátozó feltételekként az átviteli hálózatra vonatkozó feltételek szerepelnek; így a *load flow egyenletek*, valamint a generált teljesítményekre, az ágak termikus terhelésére, a csomóponti feszültségekre és a csomóponti meddőteljesítmény-kibocsátásra

vonatkozó korlátok. Általában nagyméretű, nemlineáris programozási feladat adódik. Ennek megoldására a gyakorlatban jól bevált a [17] dolgozatban közölt, a redukált gradiens és büntetőfüggvényes módszerek kombinálásából adódó algoritmus. A téma irodalmának 1975-ig terjedő feldolgozása megtalálható a [26] dolgozatban. A témakörben azóta is intenzív kutatás folyik, az Olvasó további utalásokat talál erre vonatkozóan a [20] kötetben.

A „Villamosenergia-termelés menetrendje” témakörben olyan hosszabb (egynapos, legfeljebb egyhetes) időszakot tekintünk, melyet rövidebb periódusokra bontva, a periódusokban a teljesítményigény állandónak tekinthető és megfelelő pontossággal előre becsülhető. A feladat: az erőművi egységek rendelkezésre állásának és a generált teljesítményeknek oly módon történő meghatározása, hogy a tüzelőanyagköltség és az egységek leállításából, illetve újraindításából származó hőveszteség minimális legyen. Korlátozó feltételekként leggyakrabban a teljesítmény-egyensúlyt kifejező egyenlőséget, a rendelkezésre állásokat és a termelt mennyiségeket kifejező változókat összekapcsoló feltételeket, az energiatermelésbe bevonható tüzelőanyagra vonatkozó korlátozásokat és más technológiai előírásokat leíró feltételeket vesznek figyelembe. Általában nagyméretű, vegyes, egészértékű — esetleg tiszta lineáris — programozási feladatra vezető modellt készítenek pl.: [25], [30] és [44].

A feladat megoldására alkalmazott módszerek a következők:

- kötött sorrend kijelölése az egységek be-, illetve kikapcsolására [25],
- dinamikus programozás alkalmazása [50], [70],
- egészértékű és vegyes, egészértékű programozási módszerek alkalmazása (így pl. *branch and bound módszer*) [6], [28].

A jelenlegi fejlődés fő iránya: általános modellek kidolgozása, melyekben a fenti két feladat együttes megoldására tesznek kísérletet a hálózati feltételek linearizált változatainak figyelembevételével [20].

2. Villamosenergia-rendszerek irányítása

2.1. Villamosenergia-rendszerek általános jellemzői

A villamosenergia-rendszer energiatermelő egységek, vezetékek, transzformátorok és nemzetközi kooperációs kapcsolatok olyan együttese, mely közös igazgatás, illetve felügyelet mellett látja el a fogyasztókat villamos energiával.

A rendszer kialakítása és a távlati fogyasztói igényeknek megfelelő továbbfejlesztése a rendszertervezés feladata. Az viszont az *üzemirányítás* feladata, hogy a már meglévő erőművi és hálózati berendezéseket a rendszer egésze szempontjából optimálisan üzemeltesse.

Dolgozatunk a központi üzemirányítás témaköréhez kapcsolódik. Az üzemirányítással szemben az alábbi főbb követelmények merülnek fel.

1. Ki kell elégíteni az időben változó hatásos- és reaktívteljesítmény-igényt. A hatásos teljesítménynek megfelelő energia a rendszerben nem tárolható, az ilyen igény μ sec nagyságrendű idő alatt fut át a hálózaton, és az erőművi generátoroknál jelentkezik. Szerencsére a fogyasztás összetétele olyan, hogy a teljesítményigény a perces nagyságrendben gyakorlatilag konstansnak tekinthető, ez pedig már jóval több, mint

a generátorok szabályozóinak reakcióideje. Az 1 percnél rövidebb időtartamokhoz tartozó kismértékű ingadozásokat ezek a berendezések kompenzálják. További sajátosság az, hogy a termelőnek a fogyasztásra csak kismértékű hatása van, az számára külső adottságként jelentkezik.

2. A termelt energiának műszaki jellegű minőségi követelményeket is ki kell elégténi.

- A váltakozó áram frekvenciája a névleges értéktől (50 Hz) csak a tűréshatáron belül térhet el. Ez lényeges a rendszer irányítása, ill. stabilitása szempontjából, hiszen az erőművekben szinkrongenerátorok termelik az energiát. További indokul szolgál, hogy a fogyasztók egy része (pl. szinkronmotorok, villamos órák, vasúti biztonsági berendezések) szintén névleges frekvenciára van méretezve.
- Az alaphálózati feszültségeknek az előírt határokon belül kell maradniuk ahhoz, hogy a fogyasztóknál, amelyek névleges feszültségre vannak méretezve, a feszültség ingadozása a tűréshatáron belül legyen.
- Az energiaellátásnak folyamatosnak kell lennie, sem az energiatermelő, sem a hálózati berendezésekben nem léphet fel veszélyes mértékű túlterhelés. A rendszer napi konfigurációjának olyannak kell lennie, hogy egy esetleges üzemzavar minimális kárt okozzon.
- A nemzetközi teljesítménycsere a menetrendi (szerződésben lekötött) értéktől csak tolerálható mértékben különbözhet.
- A mellékhatásként jelentkező környezetszennyezés minimális legyen.

3. A termelési költségek legyenek optimálisak.

A változó fogyasztói igényeket a fenti kényszerfeltételek mellett csak úgy lehet kielégíteni, hogy a rendszer szabályozható mennyiségeit, elsősorban az erőművek betáplálásait, az üzemirányítás folyamatosan megváltoztatja. Ezen kívül minden időszakban döntenie kell a rendszer más változtatható paramétereinek legmegfelelőbb értékeiről is, így elő kell írnia a szabályozós transzformátorok pozícióját, utasítást kell adnia a kondenzátorok, söntfajtók be- és kikapcsolására, a hálózati konfiguráció megváltoztatására. A felsorolt berendezésekkel kapcsolatban l. a 2.2.3. és az F2. részt.

A továbbiakban néhány itt vázolt kérdést részletesebben vizsgálunk.

2.1.1. A teljesítmény-egyensúly

Az üzemirányítás alapfeladata a fogyasztók mindenkori hatásos- és meddőteljesítmény-igényének kielégítése. A hálózat csomópontjainak számát N -nel jelölve (l. az F2. részt), és a hálózatot a rendszerközi vezetékek határpontjáig tekintetbe véve, a termelt és fogyasztott teljesítmények egyensúlyát az alábbi egyenletek fejezik ki:

$$(2.1) \quad \sum_{i=1}^N P_{G_i} + \sum_{i=1}^N P_{T_i} = \sum_{i=1}^N P_{D_i} + P_v$$

$$(2.2) \quad \sum_{i=1}^N Q_{G_i} + \sum_{i=1}^N Q_{T_i} + \sum_{i=1}^N Q_{K_i} + \sum_{i=1}^N Q_{C_i} - \sum_{i=1}^N Q_{L_i} = \sum_{i=1}^N Q_{D_i} + Q_v.$$

Az egyenletek közül (2.1) a hatásos-, (2.2) pedig a meddőteljesítmény-egyensúlyt fejezi ki. Az alkalmazott jelölések a következők (az i index csomópont-sorszámot jelent):

- P_{G_i}, Q_{G_i} = az erőművek hatásos, ill. meddő betáplálása;
 P_{T_i}, Q_{T_i} = a nemzetközi vezetékek határpontjain érkező összes hatásos, ill. meddő teljesítmény;
 P_{D_i}, Q_{D_i} = a fogyasztói hatásos-, ill. meddőteljesítmény-igény;
 Q_{K_i} = a szinkronkompenzátorok meddőtermelése;
 Q_{C_i} = a vezetékek kapacitásai és a söntkondenzátorok által termelt meddő teljesítmény;
 Q_{L_i} = a söntfójtók meddőfogyasztása;
 P_v, Q_v = a hálózat teljes soros hatásos-, ill. meddőteljesítmény-vesztése.

Az üzemirányítás alapfeladata a rendszer változtatható paramétereinek olyan megválasztását jelenti, hogy

- érvényes legyen a (2.1) és a (2.2) összefüggés;
- teljesüljenek a 2.1. rész bevezetésében felsorolt kényszerfeltételek;
- az $a), b)$ által szűkített halmazból olyan paraméterértékeket válasszon, melyek a rendszer egésze szempontjából „legmegfelelőbbek”;
- a változtatható mennyiségek, paraméterek adott pillanatban „legmegfelelőbb” kombinációját megvalósítsa, és a rendszert ebben az állapotában tartsa mindaddig, amíg egy új állapot felvétele nem válik indokolttá.

Az a kérdés, hogy a szabályozott mennyiségek és ezzel a rendszer állapotjellemzőinek „legjobb” kombinációja az energiarendszer milyen részletességű matematikai modelljével és milyen kritériumok alapján választandó meg, végérvényesen nem zárható le. Nyilvánvaló, hogy az ezzel kapcsolatos kutatások irányát a meglevő energiarendszerekkel szerzett tapasztalatok alapvetően befolyásolják. Jellemző például, hogy a nagy üzemzavarokat mindig a meglevő irányítási stratégia „kritikus felülvizsgálata” szokta követni.

Az összes szabályozott mennyiségnek a rendszer összes állapotjellemzőjére gyakorolt hatását figyelembe vevő globális optimalizálás — különösen, ha a szokásos kényszerfeltételeken és a gazdaságossági szempontokon túlmenően üzembiztonsági kritériumokat is érvényesít — nagyon komplex, nagyméretű nemlineáris feladat. Ebben elvben statisztikai jellegű megfontolások is szerepet kaphatnak. Az egynapos üzemirányítási feladatban azonban, ha az energiaigény előrejelzése megfelelő pontosságú, a véletlen hatásának figyelembevétele első közelítésben mellőzhető. Jelen dolgozatban kísérletet teszünk a probléma determinisztikus megfogalmazására, felépítünk egy vegyesváltozós (diszkrét-folytonos) modellt (l. 3. fejezet).

2.1.2. A jelenlegi üzemirányítási stratégia alapelemei

A jelenlegi gyakorlatban alkalmazott üzemirányítási-szabályozási eljárások a teljes feladatot egymástól többé-kevésbé függetleníthető részfeladatokra bontják, amelyek maguk is gyakran még további egyszerűsítéseket, elhanyagolásokat tartalmaznak.

A legszokásosabb szétválasztás azon a felismerésen alapul, hogy a hatásos teljesítmények (2.1) által kifejezett egyensúlya nagymértékben független a meddő teljesítményektől.

mény és a feszültség viszonyoktól (l. F4. rész). Ennek megfelelően a hatásosteljesítmény-frekvencia szabályozás különválasztható a meddőteljesítmény-feszültség szabályozástól. Az előbbi $(P-f)$ -szabályozásnak, az utóbbit pedig $(Q-V)$ -szabályozásnak is szokás nevezni. A szétválaszthatóságot nagymértékben alátámasztja az a gyakorlati körülmény, hogy az erőművekben a beavatkozási lehetőség is a fentieknek megfelelően válik ketté: a P_{Gi} termeléseket a turbina-szabályozók (primer szabályozók), a Q_{Gi} betáplálást pedig a generátorokhoz tartozó feszültség-szabályozók alapjeleinek módosításával lehet megváltoztatni.

A kétféle szabályozás közül a $(P-f)$ -szabályozás a primer feladat, minthogy a gazdasági kihatások ennél a legszembetűnőbbek, és mert a korszerű energiarendszerekben a frekvencia-eltérés tűréshatárát rendkívül szigorúan szabják meg (normális üzemben általában $\pm 0,1\%$ a tolerancia). Az erőművi frekvenciaszabályozás mechanizmusának szemléltetésére tételezzük fel, hogy a hatásosteljesítmény-igény valamely időpontban kismértékben csökken. Ekkor, mivel a turbina által leadott forgatónyomaték még a régebbi, nagyobb igénynek felel meg, a generátor forgási sebessége és azzal együtt a váltóáram frekvenciája kismértékben nő. Ezt érzékeli a turbina-szabályozó berendezés, és csökkenti a turbinába áramló gőzt. Így csökken a forgatónyomaték, és azzal a forgási sebesség is, egyúttal pedig a generátor kevesebb hatásos teljesítményt ad le, új dinamikus egyensúlyi helyzet alakul ki.

A hálózat frekvenciaszabályozása úgy történik, hogy egy vagy néhány erőműtől eltekintve (ezekben kis reakcióidejű turbinaszabályozók működnek) valamennyi erőmű a menetrend szerint táplál be aktív és reaktív teljesítményt, a frekvenciaszabályozást hatásos teljesítmény betáplálásuk révén a kiválasztott erőművek végzik.

A rendszerszintű $(Q-V)$ -szabályozásnak normális üzemben általában kisebb figyelmet szentelnek, egyrészt mert a meddőtermeléshez, ill. nyeléshez gyakorlatilag nem tartozik üzemi költség, így csak a hálózati veszteség módosításában játszik némi gazdasági szerepet, másrészt pedig, mert a hálózati feszültségek névlegestől való eltérésére legalább $\pm 5\%$, de gyakran $\pm 10\%$ is megengedhető. Rendszermentő szerepe lehet a $(Q-V)$ -szabályozásnak üzemzavarok alkalmával.

Teljesítményszabályozáson a továbbiakban mindig hatásosteljesítmény-szabályozást értünk.

2.1.3. A központi üzemirányításról

Az üzemirányítási feladat fentebb körvonalazott összetettségéből adódóan az energiarendszer központi irányításának az alábbi, idő és feladat szerinti felosztása alakult ki.

Időbeli felosztás

- Előkészítő tevékenység: a tényleges üzemvitelt megelőző, becsült terhelési adatokkal és a karbantartások figyelembevételével végzett számítások, amelyek az on-line teljesítmény- és feszültség-szabályozás alapfeltételeit határozzák meg (az egyes blokkok indításának, leállításának időpontjai; napi fejlesztési és feszültségi menetrend; export-import megállapodások; karbantartások előkészítése stb.).
- Üzemvitel: on-line (automatikus és/vagy kézi) szabályozás, amely a tényleges,

pillanatnyi adatok alapján állítja be a rendszer teljesítményegyensúlyát — ez az automatikus teljesítményszabályozás (AGC¹), amelynek részei a központi szabályozás, továbbá az erőművi blokkszabályozás —, valamint a kívánatos meddőteljesítmény-egyensúlyt — ez az automatikus vagy felügyeleti jellegű feszültség szabályozás (AVC² vagy SVC³).

Feladat szerinti felosztás

- A teljesítmény-egyensúly biztosítása: az AGC feladatban P_{G_i} meghatározása oly módon, hogy a fogyasztói igényeket a menetrendi frekvencián kielégítsék, a csereteljesítmény pedig megegyezék a menetrendben rögzített értékével.
- Döntéshozatal: az AGC feladatban a P_{G_i} összerhelést osztjuk szét a rendelkezésre álló erőművek között, azaz az egyes erőművek által termelendő hatásos teljesítményt határozzuk meg. Az AVC (SVC) feladatban a meddőforrások betáplálásainak, ill. feszültségeinek, valamint a szabályozós transzformátorok pozícióinak, ill. egyes hálózati elemek ki-bekapcsolási állapotának a meghatározása a cél. Ezt oly módon kell tennünk, hogy a fogyasztói igényeket a megengedett sínfeszültségek fennállása mellett kielégítsük, és eközben Q_{T_i} értéke is elfogadható legyen.
- Végrehajtás: az AGC feladatban az egyes erőművek által a kiszabott teljesítményváltoztatások teljesítése, amely az erőművi szabályozó, majd a blokkszabályozó segítségével zajlik le.
- Az AVC (SVC) feladatban a kiadott meddő-betáplálási, sínfeszültség módosítási, transzformátor áttétel változtatási, ill. kapcsolási parancs teljesítése, amely az erőművi meddőszabályozók, majd az erőművi generátorok, valamint a szinkronkompenzátorok gerjesztésszabályozóival, a söntfojtótekercek és kondenzátorok (nagyfeszültségű vezetékek, kábelek) megszakítóinak működtetése útján, ill. a transzformátor szabályozó léptetésével történik.

2.2. A napi ütemezési feladat megfogalmazása

A napi ütemezési feladat megoldása, vagyis az úgynevezett „napi menetrend” megadása azt jelenti, hogy a nap minden szakaszára (melyek egy óra vagy fél óra időtartamúak)

- megállapítjuk valamennyi erőműre, hogy azt melyik üzemmódjában és milyen teljesítményszinten működtetjük,
- meghatározzuk az energiaexport és import nagyságát
- és a csomóponti potenciálok értékét úgy, hogy az országot ellássuk elektromos energiával, a fogyasztói pontokon levő feszültségek a névlegestől csak csekély mértékben térjenek el, az ágak termikusan ne terhelődjenek túl, a meddőforrás csomópontok meddőteljesítmény-kibocsátása, ill. nyelése az előírt tartományba essék, figyelembe véve az erőművi, a hálózati és más gyakorlati megkorlátozásokat.

¹ Automatic Generation Control

² Automatic Voltage Control

³ Supervisory Voltage Control

Esetünkben a tervezés 25 órára vonatkozik, vagyis egy teljes napra és a következő nap első órájára. Modellünket — melyet a 3. fejezetben ismertetünk — nem egy szerűen csak valamilyen napi menetrend meghatározására dolgoztuk ki, hanem a lehetséges menetrendek közül olyanak a megkeresését tűztük ki célul, amely a figyelembe vett költségfüggvényt tekintve a lehető legkisebb költséggel jár.

A fejezet alábbi pontjaiban ennek a feladatnak egyes elemeit vizsgáljuk meg részletesebben. Így az első pontban a fogyasztást, a fogyasztói igénygörbét elemezzük, a második pontban az erőművekről ejtünk szót, a harmadik pontban az energiaelosztó villamos hálózatot ismertetjük, végül a negyedik pontban a villamos energia termelési költség összetevőit vizsgáljuk meg.

2.2.1. *A fogyasztás, a fogyasztói igénygörbe*

Az energiarendszeri szolgáltatás célja a számos, területileg szétszórt fogyasztó ellátása. Az ellátás — néhány nagyipari fogyasztótól eltekintve — az alaphálózati állomásokról (csomópontokról) táplált, kiterjedt, kisebb feszültségű elosztó hálózatok közvetítésével történik. Az alaphálózati csomópontokon fellépő teljesítményigény a nap folyamán erősen ingadozik. A vizsgálat tárgyává tett egy óra és fél óra időtartamú időszakokban mind a hatásos, mind a meddő teljesítmény ingadozása csekély, ezért állandónak tekintjük azokat.

A rendszer összes csomópontjain fellépő fogyasztások órás összegei, az erőművi önfogyasztások összege, a hálózati veszteség és az esetleges export mennyisége együttesen alkotják a pillanatnyi teljesítményigényt. Az igény időbeli változását az igény-idő függvény írja le. Ezt a függvényt a mérnöki gyakorlat röviden igénygörbének nevezi.

A napi igénygörbe az alábbi tényezőktől összetetten függ:

- évszakok (napkelte és napnyugta időpontja és a közöttük levő időszakok hossza, téli, nyári időszakmérés, fűtési szezon stb.),
- a nap típusa (munkanap, szabadnap),
- időjárási hatások (hőmérséklet, napsütés, felhőzet, szél stb.),
- TV program.

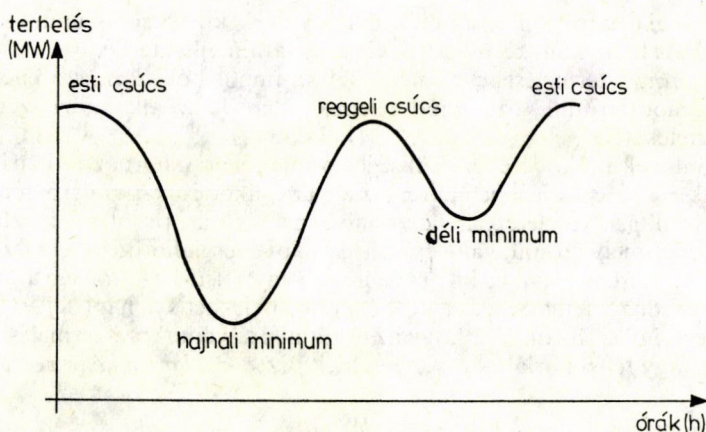
Magyarországon a napi igénygörbe igen jól — mintegy 1—2%-os eltéréssel — számítógépes programmal becsülhető előre. Ez mindig két lokális maximumot és két lokális minimumot tartalmaz.

A napi tervezés az esti csúcsterhelési időszakokkal kezdődik (mely általában a 17,00 és 19,00 óra közötti időszakra esik), továbbá 23 egyórás és 4 félórás periódusra vonatkozik.

A teljesítményigény értéke egy óra alatt csak csekély mértékben változik, ezért a számításokban az egyórás és a félórás időszakokban azt állandónak tekintjük. Ezt úgy is megfogalmazhatjuk, hogy az igénygörbét lépcsős függvénnyel helyettesítjük (lásd az 1. és 2. ábrát).

2.2.2. *Az erőművek és üzemmódjaik*

Ipari méretekben a villamos energia termelése hő-, víz- és atomerőművekben történik. Az egyes erőművekben különböző gazdaságossági és műszaki-technológiai feltételek érvényesek az ott alkalmazott energiaátalakítási módoktól, a berendezések

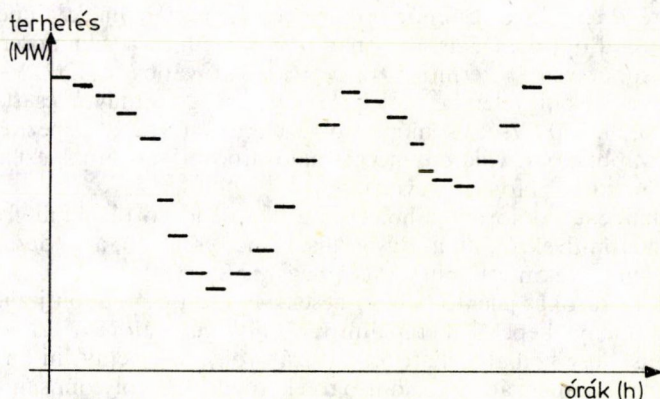


1. ábra
Jellegzetes napi igénygörbe

fajtájától, azok állapotától, a felhasznált energiahordozó minőségétől és elérésétől függően.

Ebben a dolgozatban csak hőerőműves villamosenergia-termeléssel foglalkozunk.

Az erőmű különböző berendezés-csoportjainak (blokkcsoportok, kazán-turbina-generátor együttesek) üzeme esetén az erőmű különböző *üzemmódjairól* vagy más szóval *üzemállapotairól* beszélünk. Az erőmű egyes üzemállapotai meghatározott hatásos és meddő teljesítmény tartományban üzemelnek. Egyik üzemmódból a másikba üzemmódváltással kerül az erőmű. Az *üzemmódváltás* bizonyos berendezéscsoportok leállítása, indítása útján történik. Vannak olyan üzem módok, amelyek egészen különböző berendezések működtetését jelentik. Ilyen üzem módok közötti közvetlen üzem módváltás nem megengedett a gyakorlatban, annak igen magas költsége miatt.



2. ábra
Napi igénygörbe közelítése lépcsős függvénnyel a nap egyórás és félórás szakaszokra történő felosztása esetén

Technológiai okok miatt az erőművi egységek kikapcsolás után csak legalább 4 óras állásidő eltelte után kerülhetnek vissza az áramfejlesztésbe.

Külön erőmű termeléseként fogjuk fel az importból származó energiamennyiséget. Az „importerőmű” üzemmód-jellemzői írják le az államközi szerződésekben foglalt feltételeket.

Az erőművek működéséhez szükséges primér energiahordozók ipari kooperáció útján jutnak el felhasználási helyükre. Az együttműködési kapcsolatban fellépő problémák (pl. szállítási, készletezési, vízerőműveknél tározási problémák) szükségessé tehetik, hogy valamely erőmű, vagy erőművek napi energiahordozó felhasználását korlátozzuk. Ez, matematikailag kifejezve alsó, vagy felső korlátot egyaránt jelenthet. Az energiahordozó felhasználása közvetlenül kifejezhető a napi fejlesztett energiamennyiséggel, így előfordulhat, hogy az egész napra vonatkozó termelési korlátoktól eltérő alsó, vagy felső korlátok érvényesülnek. Ezt *tüzelőanyagkényszernek* nevezzük.

2.2.3. A villamosenergia-rendszer alaphálózata

Az erőművekben generált elektromos energiának a fogyasztókhoz való szállítása a villamosenergia-hálózaton keresztül történik. A hálózat távvezeték ágai az alábbiak szerint csoportosíthatók:

- a) a 750, 400 és 220 kV-os, a nemzetközi kooperációs hálózathoz csatlakozó vezetékeknek az országhatárig terjedő szakasza,
- b) a 400 és 220 kV-os gerincvezetésekből álló országos alaphálózat,
- c) a 120 kV-os nagyfeszültségű fő-elosztóhálózat,
- d) a fentieknél kisebb feszültségű elosztóhálózat.

A hálózat különböző feszültségű részeit transzformátorokat is tartalmazó állomásokon keresztül csatlakoznak egymáshoz. A fogyasztók teljesítményigényük szerint a csomópontokon csatlakoznak a megfelelő szintű hálózathoz.

A központi irányítás számára a villamosenergia-rendszer olyan nemzetközi kapcsolódásokat is tartalmazó hálózatként jelenik meg, melyben a fogyasztások az egyes nagy feszültségű csomópontokra koncentrálnak, míg más, az előbbiekkal részben azonos csomópontokra erőművek táplálódnak. Ennek megfelelően a napi menetrend optimalizálására vonatkozó modellben a villamosenergia-hálózatnak a fenti csoportosítás szerinti a) és b) részeit, továbbá a (120 kV-os) c) hálózat azon részeit vesszük figyelembe, amelyekhez jelentős erőművek csatlakoznak, vagy amelyek a 400, ill. 220 kV-os csomópontok között sönt-ágakat képeznek, vagy képezhetnek. A továbbiakban villamosenergia-hálózaton a hálózatnak ezt a részét értjük, és ezt alaphálózatnak is fogjuk nevezni.

Az alaphálózat csomópontjaihoz erőművek, meddőforrások, ill. fogyasztók csatlakoznak. Az erőművekről, ill. a fogyasztásról már esett szó, a meddőforrásokról, ill. a nyelőkről rövid áttekintést adunk (lásd még az F2. részt).

A meddőforrásokat jelentő berendezések a hálózat csomópontjaihoz csatlakozva meddő teljesítményt képesek betáplálni (erőművi generátorok, szinkronkompenzátorok, statikus vagy szabályozható kondenzátorok), vagy elnyelni (erőművi generátorok, szinkronkompenzátorok, söntfojtótekercek). A folytonosan szabályozható források teljesítménye megadott tartományban mozoghat, a kapcsolhatóké (söntfojtók, statikus kondenzátorok) a csomóponti feszültség négyzetével arányosan változik.

A csomópontokat egymással ágak kapcsolják össze, melyek vagy távvezetékek, vagy kábelek, vagy pedig transzformátorok. Az ágak a végpontjaikra nézve villamosan szimmetrikusak, véges ellenállásuk, induktivitásuk, kapacitásuk, valamint termikus terhelhetőségük van. Villamosan ugyan nem, de az üzemirányítás szempontjából megkülönböztetjük a rendszerközi távvezetéseket, amelyek nemzeti vagy egyéb érdekelttség szerint elkülönülő rendszereket kötnek össze. Az ágak ohmos ellenállásából következően a teljesítmény-szállítás a terheléstől és a feszültségviszonyoktól függő veszteséggel jár. Kiemeljük, hogy a villamosenergia-hálózat elemei dinamikusak abban az értelemben, hogy a hálózati konfiguráció a hálózati elemek hibája, tervezett kiiktatása stb. következtében naponta, és ezen belül is esetleg rövidebb időtartamok alatt változik.

Illusztrációképpen röviden egy napon belül két szélsőséges terhelésű időszakot tekintünk.

Az éjszakai minimális terhelésű időszakban (l. 2.2.1.) a távvezetékek, kábelek gyengén vannak terhelve, ebben a helyzetben a kapacitív karakterük dominál, azaz meddő teljesítményt termelnek. A reaktív teljesítmény túltermelés a feszültségek felugráásával jár (lásd F2. részt), ennek kompenzálásában fontos szerepet játszanak a csomópontokhoz csatlakozó meddőnyelők.

A másik szélsőséges eset a csúcsterhelésű időszak, ekkor nagy meddőteljesítmény-igény mutatkozik. Ez egyrészt abból adódik, hogy a fogyasztók jelentős része meddő teljesítményt nyel (elektromos motorok), másrészt ilyenkor a teljesítményt szállító vezetékek is erősen le vannak terhelve, és ilyenkor az induktív karakterük a domináns. Ebben a helyzetben a feszültségek leesnek, és ennek kompenzálására a meddőforrások alkalmasak.

3. A villamosenergia-termelés ütemezésének egy általános matematikai programozási modellje

Ebben a fejezetben a 2.2. bekezdésben megfogalmazott ütemezési probléma egy általános matematikai programozási modellje kerül ismertetésre arra az esetre, amikor az energiatermelő rendszer csak termikus erőműveket tartalmaz. A modellt azért nevezzük általánosnak, mert megfogalmazásakor nem alkalmazunk egyszerűsítő feltevéseket azzal a céllal, hogy a megfelelő matematikai programozási modell matematikailag és számítástechnikailag kezelhető legyen.

Az 5. fejezetben egy egyszerűsített modellt fogunk ismertetni. A két modell lényegesen különbözik az üzemállapotváltozásokra vonatkozó feltevésekben.

Az általános modellben az üzemállapotváltozások kihatásait a következőképpen összegezzük:

a) Kikapcsolásnál a kikapcsolt vagy a kikapcsolttal azonos típusú, előzetesen nem működő egységek csak a minimális (4 órás) állásidő után léphetnek üzembe.

b) Bekapcsolásnál fellép a bekapcsolt egység előzetes állásidejétől függő bekapcsolási költség. A bekapcsolásra kerülő egységek üzemben léteire vonatkozóan előzetes feltételezéseket nem teszünk, az üzemeltetési időtartamot a teljes napi optimalás határozza meg.

c) Az átkapcsolásokat, amelyek egyes egységek kikapcsolásával, míg mások egyidejű bekapcsolásával járhatnak, az előző két üzemállapotváltozás összegeként kezeljük. Ez az eset ritkán fordul elő, de a lehetőségét nem zárjuk ki.

A két modell abban is különbözik, hogy míg az általános modell tetszőleges hosszúságú ütemezési időszakokra is megfogalmazható, addig az egyszerűsített modell a 4.2. bekezdésben tett egyszerűsítő feltevés miatt egynapos (ill. 25 órás) ütemezési időszakokra vonatkozik. Az egyszerűsített modellben azok az időszakok, amelyekben az igény értéke konstansnak tekinthető, $1/2$ és 1 órásak. Az általános modellben ezek tetszőleges hosszúságúak lehetnek. Az időszakokra a „periódus” elnevezést használjuk.

3.1. A modell változói

A probléma megfogalmazása (lásd 2.2. bekezdés) szerint az ütemezés időtartamának minden periódusához meg kell határoznunk az erőművekben alkalmazandó üzemmódokat, az alkalmazandó üzemmód teljesítményszintjét és a szabályozható feszültségű hálózati pontok feszültségeit.

Ennek megfelelően a modell változói az egyes periódusokbeli üzemmódokat, a teljesítményszinteket és a feszültségeket jelölő vektorok. A modell leírásakor nehézséget okoz a változók nagy száma, ill. az, hogy a változókra más és más hivatkozásmód alkalmas az éppen ismertetendő tényről függően.

Általános jelöléstechnikánk az lesz, hogy a változó vektorokat és ezek komponenseit is felső indexszel látjuk el annak a periódusnak a jelölésére, amelyre a változó vonatkozik. Ha a felső indexet elhagyjuk, arra a vektorra hivatkozunk, melyet a megfelelő felső indexszel ellátott vektoroknak a periódusok sorrendjében való egymás után fűzésével kapunk. Ezen konvenció alól kivétel lesz a hálózat feszültségviszonyait tárgyaló rész, ahol a felső index elhagyásával jelölt változó vektor egy tetszőleges, de rögzített periódushoz tartozó vektort jelöl. A hálózat feszültségi viszonyainak a tárgyalása önmagában is elég bonyolult; a megértést megkönnyíti, hogy a tárgyalásban elhagyjuk a felső indexet. Ott, ahol felső index nélküli változót ilyen értelemben használunk, erre külön felhívjuk az olvasó figyelmét.

Jelölje a továbbiakban T az ütemezési időszak periódusainak a számát, a_t pedig a t -edik periódus időtartamát. Az ütemezés időtartama eszerint $\sum_{t=1}^T a_t$. Legyen K az erőművek, $M(k)$ a k -adik erőműben alkalmazható üzemmódok száma.

Az erőművek üzemmódjainak a megnevezésére az első, második stb., $M(k)$ -adik üzemmód szóhasználatot alkalmazzuk erőművenként. Az üzemmódok sorrendjére az általános modellben nem teszünk semmi megszorítást. Lásd erre vonatkozóan a 3.2.2. bekezdés végén található megjegyzést.

A továbbiakban szükséges egyéb jelöléseket legelső előfordulásuknál definiáljuk, majd összefoglalásként áttekintjük őket a 3.5. bekezdésben.

3.1.1. Üzemmód változók

Az egyes periódusokban az erőművekben alkalmazandó üzemmód megadására bevezetjük az y^t változó vektorokat ($t=1, 2, \dots, T$), amelyek dimenziója $\sum_{k=1}^K M(k)$; és az ezekből alkotott y vektort, amely $T \sum_{k=1}^K M(k)$ dimenziós.

y^t komponensei 0 vagy 1 értékűek. Definíciójukat a következőképpen adhatjuk meg. Minden erőmű minden üzemmódjához tartozik egy komponens az erőművek, azon belül a lehetséges üzemmódok egymásutánjának sorrendjében. E szerint a sorrendezés szerint y^t l -edik komponense — a $\sum_{k=1}^{i-1} M(k) < l \leq \sum_{k=1}^i M(k)$ egyenlőtlenség teljesülése esetén — az i -edik erőmű $j = l - \sum_{k=1}^{i-1} M(k)$ -adik üzemmódjához tartozik.

A komponensekre gyakran fogunk kettős indexszel is hivatkozni, ahol az első index a komponens által jellemzett erőműnek, a második index a jellemzett üzemmódnak a sorszáma.

Legyen az itt leírt l -edik komponens értéke 1, ha a t -edik periódusban az i -edik erőműben a j -edik üzemmód alkalmazandó, egyébként 0.

Mivel minden erőműben mindenkor egyetlen üzemmód üzemel, e definícióból azonnal következik, hogy teljesülnie kell a

$$(3.1) \quad \sum_{j=1}^{M(i)} y_{ij}^t = 1, \quad i = 1, 2, \dots, K, \quad t = 1, 2, \dots, T$$

összefüggéseknek.

A felső index nélküli y vektor megállapodásunk szerint az y^t , $t = 1, 2, \dots, T$ vektorok egymás után fűzésével keletkezik. Így pl. $y_i^t = y_{ij}^t$ az y vektor

$$(t_0 - 1) \sum_{k=1}^K M(k) + l = (t_0 - 1) \sum_{k=1}^K M(k) + \sum_{k=1}^{i-1} M(k) + j\text{-edik komponense.}$$

A későbbiekben szükségünk lesz a tervezési időszakot megelőző időszak legutolsó periódusában alkalmazott üzemmódok ismeretére is. Az erre vonatkozó információ megadható ezen legutolsó periódusra vonatkozó üzemmód változók értékeinek megadásával. Jelöljük y^0 -val a megfelelő üzemmód vektort. Ez a modellben $\sum_{k=1}^K M(k)$ dimenziós, 0–1 komponenseket tartalmazó konstans vektor.

Megjegyzés:

1. Az olvasóban esetleg felvetődik az a kérdés, hogy mi történik, ha két egymást követő tervezési időszakban az erőművekben alkalmazható üzemmódok nem azonosak, illetve ha egy tervezési időszakon belül változik az üzemmód készlet. (Ez előfordulhat, hiszen például valamilyen karbantartási munka a nap bármely időpontjában befejeződhet, s így addig nem alkalmazható gépi egységek is működtethetők ezt követően.)

A kérdés megválaszolása egyszerű: a tervezési időszakban alkalmazható üzemmódokat úgy adjuk meg, hogy ezek halmaza a lehető legbővebb legyen. Külön üzemmódnak tekintjük azt az üzemmódot is, amelynek alkalmazása csak néhány periódusban megengedett. Ezekre vonatkozóan megköveteljük, hogy azokban a periódusokban, amelyekben ez az üzemmód nem alkalmazható, a megfelelő üzemmód változó értéke 0 legyen.

Ez a megállapodás lehetővé teszi, hogy az $M(k)$, $k=1, 2, \dots, K$ üzemmód készlet független legyen az időtől, és egy-egy üzemmód sorszáma mindig ugyanazt az üzemmódot jelölje.

2. Felvetődhet az a kérdés is, hogy mi történik, ha valamely erőműben csak egyetlen üzemmód alkalmazható, azaz valamely $M(i)=1$. Ebben az esetben a modellben ezen üzemmód alkalmazásának teljesítményszintjét kell csak megadni, a megfelelő üzemmód változó használata felesleges, hiszen ennek értéke csak 1 lehet.

Ennek ellenére, a modell leírását megkönnyítendő, ezekkel az erőművekkel, ill. üzemmódokkal nem foglalkozunk külön, hanem használjuk az y_{it} változót, s a (3.1) egyenlőség megkövetelésével biztosítjuk az $y_{it}=1$ teljesülését.

3. Megjegyezzük még, hogy az egyszerűsített modell üzemmód változóit az itenitől eltérő módon definiáljuk.

3.1.2. Teljesítmény változók

Az egyes periódusokban az erőművekben alkalmazandó üzemmódok teljesítményszintjének megadására a \mathbf{p}^t teljesítmény-változó vektorokat használjuk. Ezek dimenziója $\sum_{k=1}^K M(k)$. Összefűzésükkel ($t=1, 2, \dots, T$) keletkezik a \mathbf{p} vektor, amely $T \sum_{k=1}^K M(k)$ dimenziós.

\mathbf{p}^t definíciója a következő:

Minden erőmű minden üzemmódjához hozzárendeljük \mathbf{p}^t egy komponensét (az erőművek, azon belül az üzemmódok sorrendjében, ugyanúgy, ahogyan az üzemmód változóknál tettük). A komponensekre kettős indexszel is fogunk hivatkozni, ahol az első index a megfelelő erőmű, a második az üzemmód sorszáma. Ezek szerint a p_{ij}^t komponens a \mathbf{p}^t vektor $\sum_{k=1}^{i-1} M(k) + j$ -edik komponense, ill. a \mathbf{p} vektor definíciója szerint ennek $(t-1) \sum_{k=1}^K M(k) + \sum_{k=1}^{i-1} M(k) + j$ -edik komponense.

Minden erőmű minden üzemmódjához ismert a működtetésnél megengedett minimális és maximális teljesítményszint. Jelölje ezeket P_{ij}^{\min} és P_{ij}^{\max} , $i=1, 2, \dots, K$, $j=1, 2, \dots, M(i)$.

A p_{ij}^t komponens értékét definiáljuk a következőképpen: legyen a $p_{ij}^t=0$, ha a t -edik periódusban, az i -edik erőműben nem a j -edik üzemmód kerül alkalmazásra, egyébként az i -edik erőmű j -edik üzemmódjában a működési szintnek a minimális érték feletti része.

Ezzel a definícióval az i -edik erőmű j -edik üzemmódjának teljesítményszintje a

$$P_{ij}^{\min} y_{ij}^t + p_{ij}^t$$

összeggel adható meg, és teljesülnie kell a

$$(3.2) \quad P_{ij}^{\min} y_{ij}^t \leq P_{ij}^{\min} y_{ij}^t + p_{ij}^t \leq P_{ij}^{\max} y_{ij}^t, \quad \text{illetve:} \quad 0 \leq p_{ij}^t \leq (P_{ij}^{\max} - P_{ij}^{\min}) y_{ij}^t$$

összefüggésnek. A fenti definíció értelmében a p_{ij}^t komponens csak akkor vehet fel nem zéró értéket, ha $y_{ij}^t = 1$. E követelmény teljesülését a (3.2) egyenlőtlenség jobb oldalán szereplő $P_{ij}^{\max} y_{ij}^t$ szorzatban az y_{ij}^t tényező biztosítja. (Nem igaz azonban, hogy $y_{ij}^t = 1$ teljesüléséből következik az, hogy $p_{ij}^t > 0$, hiszen ha az üzemmód a megengedett minimális szinten működik, akkor $p_{ij}^t = 0$.)

3.1.3. Feszültségi változók

A modellben változóknak tekintjük az átviteli hálózat valamennyi csomópontja komplex feszültségének valós és képzetes részét. Jelöljük a t -edik periódusban a csomóponti feszültségek valós részét v_1^t, \dots, v_N^t -vel, a képzetes részét pedig w_1^t, \dots, w_N^t -vel, ahol N a hálózat csomópontjainak száma. A megfelelő vektorokat jelölje \mathbf{v}^t , ill. \mathbf{w}^t . A továbbiakban, amennyiben egy rögzített periódust tekintünk, a felső indexeket el fogjuk hagyni.

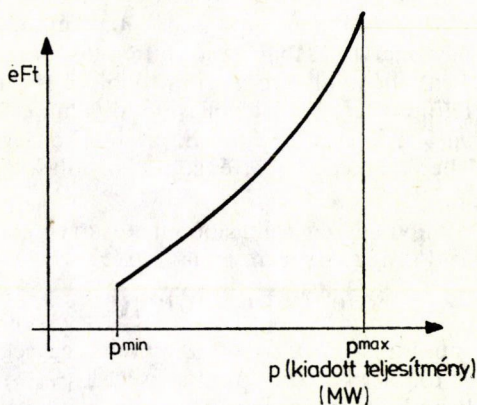
3.2. A célfüggvény

A modell minimalizálandó célfüggvényét a villamosenergia-termelés költsége szolgáltatja. Ez tartalmazza az erőművi blokkok adott szinten való működtetéséhez szükséges tüzelőanyag költségét, az erőművi blokkok bekapcsolásakor fellépő, állásukból származó hővesztési és az üzemmódváltással kapcsolatos állagromlási költséget, valamint a hálózati veszteségből eredő költséget. E három összetevőt külön-külön tárgyaljuk.

3.2.1. Erőművi blokkok termelési költsége

Az erőművi blokkok üzemeléséből származó költségrészt a következőképpen adhatjuk meg:

Minden erőmű minden lehetséges üzemmódjához tartozik egy arra jellemző költség-görbe (lásd 3. ábra).



3. ábra

Jellegzetes fejlesztési költség-görbe

Legyen $f_{ij}(P)$ az i -edik erőmű j -edik üzemmódjához tartozó költséggörbét megadó függvény, ahol a P független változó az üzemeltetés teljesítményszintje, a függvényérték pedig a P szintű termelés esetén időegység alatt elhasznált tüzelőanyag ára.

$f_{ij}(P)$ értelmezési tartománya a $[P_{ij}^{\min}, P_{ij}^{\max}]$ intervallum, hiszen az i -edik erőmű j -edik üzemmódja csak ezen teljesítménykorlátok között működhet. Az i -edik erőmű j -edik üzemmódja működésből származó költség rész nyilvánvalóan 0, ha az erőműben nem ez az üzemmód működik.

Vezessük be a következő jelölést:

Legyen $K_{ij} = f_{ij}(P_{ij}^{\min})$, vagyis az i -edik erőmű j -edik üzemmódjának működése esetén a minimális teljesítmény kibocsátásának költsége.

Legyen $k_{ij}(P) = f_{ij}(P_{ij}^{\min} + P) - f_{ij}(P_{ij}^{\min})$, azaz a $P_{ij}^{\min} + P$ teljesítményszinten való működtetés esetén, a minimális P_{ij}^{\min} mennyiségen felül termelt P energiamennyiség többlet termelési költsége. A $k_{ij}(P)$ függvény értelmezési tartománya a $[0, P_{ij}^{\max} - P_{ij}^{\min}]$ intervallum.

Megállapodásunk szerint a t -edik periódus időtartamát a_t jelöli. Az eddigiek alapján az i -edik erőmű j -edik üzemmódjának t -edik periódusbeli üzemeléséből származó költség a következő:

$$(3.3) \quad a_t \{K_{ij} y_{ij}^t + k_{ij}(p_{ij}^t)\}.$$

Így a teljes tervezési intervallumban az erőművi blokkok működéséből származó költség rész:

$$(3.4) \quad \sum_{t=1}^T a_t \sum_{i=1}^K \sum_{j=1}^{M(i)} \{K_{ij} y_{ij}^t + k_{ij}(p_{ij}^t)\}.$$

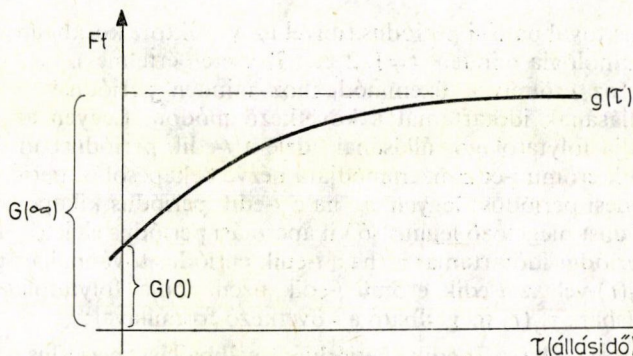
3.2.2. Berendezések állásából, újraindításából származó költség: átállási költség

Az üzemmódok állásából származó költség figyelembevétele valójában annak a költségnek a figyelembevételét jelenti, amely az egyes erőművi egységek bekapcsolásánál, az erőművi egység egy bizonyos állásidő utáni újraindításakor jelentkezik. Egy-egy erőművi egység vonatkozásában ez a költség tartalmaz egy, a berendezések kapcsolása következtében fellépő állagromlási és többlet-karbantartási költséget, valamint a működéshez felfűtött testeknek az állásidő folyamán történő lehűlése következtében szükségessé váló felfűtési költségét. Ez utóbbi költség rész — mivel a hűlés során exponenciális jellegű hőleadás történik — az állásidő exponenciális függvénye.

A 4. ábrán egy erőművi egység τ állásidő eltelte utáni újraindításának költség-görbéje látható, az állásidő függvényében. Ez a költség-görbe a

$$(3.5) \quad g(\tau) = G(0) + \{G(\infty) - G(0)\}(1 - e^{-c\tau})$$

képlettel adható meg, ahol a c , $G(0)$, $G(\infty)$ az erőművi egységre jellemző állandók, $G(0)$ az állagromlást és többlet karbantartási költséget, $\{G(\infty) - G(0)\}(1 - e^{-c\tau})$ pedig a hőveszteség miatt fellépő költség részét jelenti. $G(\infty)$ a hosszú állásidő utáni ún. hidegen indítás költsége; $c > 0$, mivel a hőveszteség-költség az állásidő növekedésével együtt növekszik. A $g(\tau)$ függvény értelmezve van minden $\tau \geq 0$ esetén, annak ellené-



4. ábra

Erőművi egységek leállítás-újraindítási (állási) költségeinek alakulása az állásidő függvényében

re, hogy az erőművi egységek vagy folyamatosan működnek, és ekkor természetesen nincs állás-újraindítási költség, vagy a tényleges állásidő értéke legalább 4 óra. A nagytömegű berendezések tranziens hőigénybevételét korlátozó ajánlások, szabványok, garanciák, üzemviteli előírások miatt ugyanis a leállított berendezések csak legalább 4 óra elteltével indíthatók újra.

Hogy az erőművi egységek állásából származó költséget a modellben — amelyben üzemmódok működéséről, ill. állásáról történik döntés — figyelembe vehessük, tisztáznunk kell az erőművi egységek és az üzemmódok kapcsolatát. Ha minden üzemmód más és más erőművi egység működését jelentené, akkor az üzemmód állásának ideje és egy erőművi egység állásának ideje azonos lenne, és az erőművi egységek állásköltségeinek számításához az üzemmódok állásidejét kellene figyelembe vennünk, amely az y_{ij}^t üzemmód változók függvényeként megadható. Az erőművek üzemmódjai és az erőművi egységek között azonban nincs ilyen kölcsönösen egyértelmű kapcsolat. Az erőművek üzemmódjai az erőművi egységek különböző, műszakilag lehetséges együttműködéseként adódnak. Példaként tekintsük a következőt:

Tegyük fel, hogy valamely erőműben három erőművi egység van, jelölje ezeket [1], [2], [3]. Amennyiben ennek az erőműnek két lehetséges üzemmódja [A] és [B] lenne, ahol [A] az [1] erőművi egység működését, [B] pedig a [2] és [3] egységek működését jelentené, akkor az erőművi egységek állásideje azonos lenne az üzemmódok állásidejével, és az állásköltség egyszerűen számítható lenne. Ha azonban lenne egy harmadik, [C] üzemmód is, amely mindhárom erőművi egység működését jelentené, akkor az erőművi egységek állásideje már nem lenne azonos az üzemmódok állásidejével.

A következőkben az erőművi egységek állásköltségének figyelembevételére adunk egy módszert. Ennek leírásában a következő terminológiát és jelöléseket használjuk:

Az i -edik erőmű j -edik üzemmódja vonatkozásában az egyes periódusok kezdetének időpontja lehet *kikapcsolási* vagy *bekapcsolási* időpont, ill. változatlan működés esetén a periódus *folytatólagos működési* vagy *továbbállási periódus*. Így a t -edik periódus kezdetének időpontja kikapcsolási időpont, ha $y_{ij}^{t-1}=1$ és $y_{ij}^t=0$. Ha $y_{ij}^{t-1}=0$ és $y_{ij}^t=1$, akkor ez az időpont bekapcsolási időpont. A t -edik periódus folytatólagos működési periódus, ha $y_{ij}^{t-1}=1$ és $y_{ij}^t=1$. Egyébként $y_{ij}^{t-1}=0$ és $y_{ij}^t=0$

esetén a periódus továbbállási periódus (mivel az y^0 vektort korábban már definiáltuk, a fenti terminológia minden $t=1, 2, \dots, T$ esetén értelmes).

Definiáljuk az erőművek üzemmódjaihoz minden periódusban az üzemmód folytatólagos állásának időtartamát a következő módon. Legyen az i -edik erőmű j -edik üzemmódja folytatólagos állásának ideje a t -edik periódusban 0, ha a t -edik periódus az i -edik erőmű j -edik üzemmódjára nézve bekapcsolási periódus vagy folytatólagos működési periódus; legyen a_t , ha a t -edik periódus kikapcsolási periódus; és a t -edik periódust megelőző legutolsó kikapcsolási periódus elejétől eltelt idő, beleértve a t -edik periódus időtartamát is, ha a t -edik periódus továbbállási periódus.

Jelöljük $\tau_{ij}(t)$ -vel az i -edik erőmű j -edik üzemmódja folytatólagos állásidejét a t -edik periódusban. $\tau_{ij}(t)$ megadható a következő formulával:

$\tau_{ij}(t) = \tau_{ij}(t-1) + a_t$, ha a t -edik periódus továbbállási periódus, azaz $y_{ij}^{t-1} = 0$ és $y_{ij}^t = 0$;

$\tau_{ij}(t) = a_t$, ha a t -edik periódus kezdetének időpontja kikapcsolási időpont, azaz $y_{ij}^{t-1} = 1$ és $y_{ij}^t = 0$;

$\tau_{ij}(t) = 0$, ha a t -edik periódus folytatólagos működési periódus vagy kezdetének időpontja bekapcsolási időpont, azaz vagy $y_{ij}^{t-1} = 1$ és $y_{ij}^t = 1$, vagy $y_{ij}^{t-1} = 0$ és $y_{ij}^t = 1$.

$\tau_{ij}(t)$ -nek az üzemmódváltozó-komponensektől való függése megadható a következő szorzat formájában is:

$$(3.6) \quad \tau_{ij}(t) = \{\tau_{ij}(t-1) + a_t\} (1 - y_{ij}^t),$$

hiszen $y_{ij}^t = 1$ esetén e szorzat értéke 0, egyébként $y_{ij}^t = 0$ és $y_{ij}^{t-1} = 1$ esetén $\tau_{ij}(t) = a_t$, hiszen $\tau_{ij}(t-1) = 0$. Végül $y_{ij}^t = 0$ és $y_{ij}^{t-1} = 0$ esetén $\tau_{ij}(t) = \tau_{ij}(t-1) + a_t$.

A definíció teljességéhez meg kell még adni $\tau_{ij}(0)$ értelmezését. Legyen $y_{ij}^0 = 1$ esetén (azaz ha az üzemmód az előző tervezési időszak befejezésekor működik) $\tau_{ij}(0) = 0$, és $y_{ij}^0 = 0$ esetén legyen $\tau_{ij}(0)$ a tervezési időszakot megelőzően eltelt állás-idő, vagyis $\tau_{ij}(0)$ az előző tervezési időszakra vonatkozó $\tau_{ij}(T)$.

Az üzemmódokhoz hasonlóan erőművi egységek vonatkozásában is használjuk a kikapcsolási, ill. bekapcsolási időpont, folytatólagos működési, ill. továbbállási periódus elnevezéseket, és definiáljuk erőművi egységek folytatólagos állásának az idejét. Ehhez használjuk a következő jelöléseket.

Legyen $N(i)$ az i -edik erőmű egységeinek száma. Hivatkozzunk az egységekre sorszámokkal, a sorszámokból alkotott indexhalmaz legyen $L(i) = \{1, 2, \dots, N(i)\}$. Jelöljük $J(i, j)$ -vel az $L(i)$ azon részhalmazát, amelynek megfelelő egységek együttes működése az i -edik erőmű j -edik üzemmódját szolgáltatja ($i=1, 2, \dots, K$; $j=1, 2, \dots, M(i)$).

Az i -edik erőmű k_0 sorszámú egysége működik a t -edik periódusban, ha $y_{ij_0}^t = 1$ esetén $k_0 \in J(i, j_0)$. Abban az esetben, ha $y_{ij_0}^t = 1$ és $k_0 \notin J(i, j_0)$, a k_0 sorszámú egység a t -edik periódusban áll.

A fenti információ $\sum_{j: k_0 \in J(i, j)} y_{ij}^t$ összeg értékével is megadható, ahol az összegzés azon j értékekre történik, amelyekre $k_0 \in J(i, j)$.

$$\sum_{j: k_0 \in J(i, j)} y_{ij}^t = \begin{cases} 1, & \text{ha az } i\text{-edik erőmű } k_0 \text{ egysége} \\ & \text{működik a } t\text{-edik periódusban,} \\ 0, & \text{egyébként.} \end{cases}$$

A t -edik periódus kezdetének időpontja az i -edik erőmű k_0 egységére kikapcsolási időpont, ha

$$\sum_{j: k_0 \in J(i, j)} y_{ij}^{t-1} = 1 \quad \text{és} \quad \sum_{j: k_0 \in J(i, j)} y_{ij}^t = 0.$$

A bekapcsolási időpont, folytatólagos működési, ill. továbbállási periódus elnevezést hasonlóan definiáljuk — mindenütt a $\sum_{j: k_0 \in J(i, j)} y_{ij}^t$ összegek figyelembevételével.

Szükségünk lesz az egyes erőművi egységek folytatólagos állásának idejére, — ezt is az üzemmódok folytatólagos állásának idejével analóg módon definiáljuk.

Jelölje $\xi_{ik}(t)$ az i -edik erőmű k -adik egységének folytatólagos állásidejét a t -edik periódusban. $\xi_{ik}(t)$ megadható a következő formulával:

$$\begin{aligned} \xi_{ik}(t) &= \xi_{ik}(t-1) + a_t, & \text{ha} & \quad \sum_{j: k \in J(i, j)} y_{ij}^{t-1} = 0 \quad \text{és} \quad \sum_{j: k \in J(i, j)} y_{ij}^t = 0, \\ \xi_{ik}(t) &= a_t, & \text{ha} & \quad \sum_{j: k \in J(i, j)} y_{ij}^{t-1} = 1 \quad \text{és} \quad \sum_{j: k \in J(i, j)} y_{ij}^t = 0, \\ \xi_{ik}(t) &= 0, & \text{ha} & \quad \sum_{j: k \in J(i, j)} y_{ij}^{t-1} = 1 \quad \text{és} \quad \sum_{j: k \in J(i, j)} y_{ij}^t = 1, \\ & & \text{vagy} & \quad \sum_{j: k \in J(i, j)} y_{ij}^{t-1} = 0 \quad \text{és} \quad \sum_{j: k \in J(i, j)} y_{ij}^t = 1, \end{aligned}$$

ill. megadható a következő szorzat formájában is:

$$\xi_{ik}(t) = \{\xi_{ik}(t-1) + a_t\} \left(1 - \sum_{j: k \in J(i, j)} y_{ij}^t\right).$$

A definíció teljességéhez szükséges $\xi_{ik}(0)$ értékét $\tau_{ij}(0)$ -val analóg módon definiáljuk.

Az üzemmódok állásidejének és az egységek állásidejének definíciójából következik, hogy

$$\xi_{ik}(t) = \min_{j: k \in J(i, j)} \tau_{ij}(t), \quad k = 1, 2, \dots, N(i), \quad i = 1, 2, \dots, K,$$

és ha valamely egység működése csak egyetlen üzemmód működése esetén szükséges, akkor ezen egység és üzemmód állásideje megegyezik.

Az energiatermelő rendszer egészében fellépő állásköltséget periódusonként, erőművenként, ill. erőművi egységenként vesszük figyelembe, majd összegezzük az így adódó költség részeket. Minden periódusban az erőművi egységnek az adott periódusbeli szerepéből származó költség részt számítjuk fel.

Jelölje $g_{ik}(\tau)$ az i -edik erőmű k -adik erőművi egysége állásköltség függvényét, $k=1, 2, \dots, N(i)$, $i=1, 2, \dots, K$.

Vizsgáljuk a $g_{ik}(\xi_{ik}(t))$ függvényértéket. Ha az i -edik erőmű k -adik erőművi egysége a t -edik periódusban nem működik, akkor ez a függvényérték a t -edik periódus végéig eltelt állásidőhöz tartozó állásköltség. Ha az i -edik erőmű k -adik erőművi egysége működik a t -edik periódusban, akkor $\xi_{ik}(t)=0$, és a megfelelő függvényérték az állagromlás költségének felel meg.

Így az i -edik erőmű k -adik erőművi egységének t -edik periódusbeli állapotából származó költség a $g_{ik}(\tau)$ függvény felhasználásával a következő formulával adható

meg:

$$(3.7) \quad \left\{1 - \left(\sum_{j:k \in J(i,j)} y_{ij}^{t-1}\right) \left(\sum_{j:k \in J(i,j)} y_{ij}^t\right)\right\} g_{ik}(\xi_{ik}(t)) - \left\{1 - \sum_{j:k \in J(i,j)} y_{ij}^t\right\} g_{ik}(\xi_{ik}(t-1)).$$

Ennek belátására nézzük sorra a következő négy esetet:

a) Ha az i -edik erőmű k -adik erőművi egysége a t -edik periódusban és az azt megelőző periódusban is áll, azaz:

$$\sum_{j:k \in J(i,j)} y_{ij}^{t-1} = 0 \quad \text{és} \quad \sum_{j:k \in J(i,j)} y_{ij}^t = 0,$$

akkor a (3.7) kifejezés értéke:

$$g_{ik}(\xi_{ik}(t)) - g_{ik}(\xi_{ik}(t-1)),$$

ami a t -edik periódusbeli, $\xi_{ik}(t-1)$ időponttól a $\xi_{ik}(t)$ időpontig tartó a_t időtartamú továbbállás költsége.

b) Ha az i -edik erőmű k -adik erőművi egysége vonatkozásában a t -edik periódus kezdetének időpontja kikapcsolási időpont, azaz a $t-1$ -edik periódusban olyan üzemmód működik, amelynek működéséhez a k -adik egység működése szükséges ($\sum_{j:k \in J(i,j)} y_{ij}^{t-1} = 1$), a t -edik periódusban pedig az ilyen üzemmódok egyike sem üzemel ($\sum_{j:k \in J(i,j)} y_{ij}^t = 0$), akkor (3.7) értéke

$$g_{ik}(a_t) - g_{ik}(0),$$

ami az a_t idejű állásból származó költség.

c) Ha az i -edik erőmű k -adik erőművi egysége a t -edik és az azt megelőző periódusban is működik ($\sum_{j:k \in J(i,j)} y_{ij}^{t-1} = 1$ és $\sum_{j:k \in J(i,j)} y_{ij}^t = 1$), akkor (3.7) értéke 0.

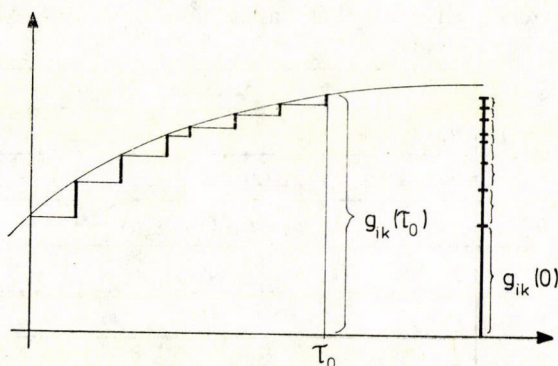
d) Ha a t -edik periódus kezdete bekapcsolási periódus az i -edik erőmű k -adik erőművi egységére nézve, — azaz: az i -edik erőmű k -adik erőművi egysége a t -edik periódusban működik, ($\sum_{j:k \in J(i,j)} y_{ij}^t = 1$), de a megelőző periódusban nem ($\sum_{j:k \in J(i,j)} y_{ij}^{t-1} = 0$), akkor (3.7) értéke $g_{ik}(0)$, ami a megelőző állási időszakhoz tartozó állagromlást kifejező érték.

A vizsgált (3.7) kifejezés értékeinek $t=1, 2, \dots, T$ -re való összegzése esetén végül is az i -edik erőmű k -adik erőművi egységének az egész időszak alatt betöltött szerepéből származó költséget az állásköltség-függvény periódushossznyi függetlenváltozó növekedéshez tartozó függvényérték növekményeinek összegeként állítjuk elő. A $g_{ik}(0)$ állagromlás-költséget az egység bekapcsolásának időpontjában vesszük figyelembe (lásd 5. ábra). Folytonos működési periódusokban állásköltség nem lép fel.

Illusztrációként tekintsünk egy konkrét esetet:

Tegyük fel, hogy az ötödik erőműben három erőművi egység van, amelyek közül az 1. és 3. együttműködése az 1., mindhárom egység együttműködése a 2. üzemmódot szolgáltatja. $M(5)=2$, $N(5)=3$, $J(5, 1)=\{1, 3\}$, $J(5, 2)=\{1, 2, 3\}$.

Tegyük fel, hogy a tervezési időszakot megelőzően és annak első három periódusában az első üzemmód működik, ezt követően 4 periódusban a második üzemmód működik. Végül a tervezési időszak végéig az első üzemmód üzemel. Legyen a terve-



5. ábra

A τ_0 állásidőhöz tartozó átállási költség előállítás függvényérték-növekmények és $g_{ik}(0)$ összegeként

zési időszak 10 periódus hosszúságú. Tegyük fel, hogy a második üzemmód kikapcsolása a tervezési időszak előtt 2 órával történt. Az üzemmód változók értéke ekkor a következő:

$$\begin{aligned} y_{51}^0 &= 1, & y_{51}^1 &= 1, & y_{51}^2 &= 1, & y_{51}^3 &= 1, & y_{51}^4 &= 0, & y_{51}^5 &= 0, & y_{51}^6 &= 0, & y_{51}^7 &= 0, \\ & & y_{51}^8 &= 1, & y_{51}^9 &= 1, & y_{51}^{10} &= 1; \\ y_{52}^0 &= 0, & y_{52}^1 &= 0, & y_{52}^2 &= 0, & y_{52}^3 &= 0, & y_{52}^4 &= 1, & y_{52}^5 &= 1, & y_{52}^6 &= 1, & y_{52}^7 &= 1, \\ & & y_{52}^8 &= 0, & y_{52}^9 &= 0, & y_{52}^{10} &= 0. \end{aligned}$$

Az üzemmódok állásideje $\tau_{51}(t)$ és $\tau_{52}(t)$, $t=0, 1, 2, \dots, 10$ a következő:

$$\begin{aligned} \tau_{51}(0) &= 0, & \tau_{51}(1) &= \tau_{51}(2) = \tau_{51}(3) = 0, & \tau_{51}(4) &= a_4, & \tau_{51}(5) &= a_4 + a_5, \\ \tau_{51}(6) &= a_4 + a_5 + a_6, & \tau_{51}(7) &= a_4 + a_5 + a_6 + a_7, & \tau_{51}(8) &= 0, & \tau_{51}(9) &= 0, \\ & & \tau_{51}(10) &= 0, \\ \tau_{52}(0) &= 2, & \tau_{52}(1) &= 2 + a_1, & \tau_{52}(2) &= 2 + a_1 + a_2, & \tau_{52}(3) &= 2 + a_1 + a_2 + a_3, \\ \tau_{52}(4) &= \tau_{52}(5) = \tau_{52}(6) = \tau_{52}(7) = 0, & \tau_{52}(8) &= a_8, & \tau_{52}(9) &= a_8 + a_9, \\ & & \tau_{52}(10) &= a_8 + a_9 + a_{10}. \end{aligned}$$

Az erőművi egységek állásideje $\xi_{51}(t)$, $\xi_{52}(t)$ és $\xi_{53}(t)$ a következő ($t=0, 1, 2, \dots, 10$):

$$\begin{aligned} \xi_{51}(t) &= \min_{j: 1 \in J(t, j)} \tau_{5, j}(t) = \begin{cases} \tau_{51}(t) = 0, & t = 0, 1, 2, 3, 8, 9, 10 \\ \tau_{52}(t) = 0, & t = 4, 5, 6, 7; \end{cases} \\ \xi_{52}(t) &= \min_{j: 2 \in J(5, j)} \tau_{5, j}(t) = \tau_{52}(t); \\ \xi_{53}(t) &= \min_{j: 3 \in J(5, j)} \tau_{5, j}(t) = \min(\tau_{51}(t), \tau_{52}(t)) = 0. \end{aligned}$$

Így az erőművi egységek periódusonkénti állapotából származó költségek:

periódus	1. egység	2. egység	3. egység
1.	0	$g_{52}(2+a_1)-g_{52}(2)$	0
2.	0	$g_{52}(2+a_1+a_2)-g_{52}(2+a_1)$	0
3.	0	$g_{52}(2+a_1+a_2+a_3)-g_{52}(2+a_1+a_2)$	0
4.	0	$g_{52}(0)$	0
5.	0	0	0
6.	0	0	0
7.	0	0	0
8.	0	$g_{52}(a_8)-g_{52}(0)$	0
9.	0	$g_{52}(a_8+a_9)-g_{52}(a_8)$	0
10.	0	$g_{52}(a_8+a_9+a_{10})-g_{52}(a_8+a_9)$	0

A periódusonként figyelembe vett állásköltségek összege:

$$g_{52}(2+a_1+a_2+a_3)-g_{52}(2)+g_{52}(0)+g_{52}(a_8+a_9+a_{10})-g_{52}(0).$$

Ez megfelel annak, hogy az erőműben a 2. erőművi egység a tervezési időszakot megelőző 2 órás állás után három perióduson át áll, $(g_{52}(2+a_1+a_2+a_3)-g_{52}(2))$, ezután bekapcsolásra kerül $(g_{52}(0))$, végül a nyolcadiktól a tizedik periódusig áll $(g_{52}(a_8+a_9+a_{10})-g_{52}(0))$.

Fentiek alapján a teljes energiarendszer állásköltsége

$$(3.8) \quad \sum_{t=1}^T \sum_{i=1}^K \sum_{k=1}^{N(i)} \left[\left\{ 1 - \left(\sum_{j:k \in J(i,j)} y_{ij}^{t-1} \right) \left(\sum_{j:k \in J(i,j)} y_{ij}^t \right) \right\} g_{ik}(\xi_{ik}(t)) - \left\{ 1 - \sum_{j:k \in J(i,j)} y_{ij}^t \right\} g_{ik}(\xi_{ik}(t-1)) \right].$$

Ez az összeg nem tartalmazza a vizsgált időszak utolsó periódusában álló erőművi egységek állagromlás költségét. Ezt a költség részt ugyanis az újraindítás időpontjában vesszük figyelembe, újraindítás pedig itt nem következik be. Az utolsó periódusban álló erőművi egységek állagromlás költségének összege:

$$(3.9) \quad \sum_{i=1}^K \sum_{k=1}^{N(i)} g_{ik}(0) \left(1 - \sum_{j:k \in J(i,j)} y_{ij}^T \right)$$

a következő tervezési időszakban lép fel.

Az állásköltség számolásakor nem foglalkozunk azzal, hogy a tényleges állásidő leállított berendezések esetén mindig legalább négy óra. Ennek teljesüléséről a modell feltételrendszerének megfogalmazásánál gondoskodunk.

Megjegyzés:

Az erőművi egységek állásával, újraindításával kapcsolatosan fellépő költség figyelembevételének módja általánosabb mint ami az erőművek üzemviteli gyakorlatának ismeretében szükséges. Igaz ugyan, hogy egy adott erőmű egységei sokféle módon működhetnek együtt, vagyis az erőművekben sokféle üzemmód lehetséges, s ezek megadhatók a megfelelő erőművi egységek sorszámaiból álló indexhalmazokkal (ezeket $J(i, j)$, $j=1, 2, \dots, M(i)$ -vel jelöltük). Az üzemviteli gyakorlat szerint azonban egy adott tervezési időszakban az összes lehetséges üzemmódoknak csak egy része megengedett, mégpedig csak olyan része, mely sorbarendeázhető úgy, hogy a megfelelő $J(i, j)$ halmazokra

$$J(i, j_1) \supset J(i, j_2) \supset \dots \supset J(i, j_r)$$

teljesül. Ritkán, de előfordul egy ennél általánosabb eset, amikor a tervezési időszakban megengedett üzemmódok úgy rendezhetők csak sorba, hogy a megfelelő $J(i, j)$ halmazok két fenti tulajdonságú csoportra bonthatók. Legyenek ezek

$$J(i, j_1) \supset J(i, j_2) \supset \dots \supset J(i, j_r),$$

$$J(i, l_1) \supset J(i, l_2) \supset \dots \supset J(i, l_s).$$

A két csoport általában nem független, $J(i, l_1)$ része valamely, az első csoportba tartozó indexhalmaznak, s valamely az első csoportba tartozó halmaz része $J(i, l_s)$ -nek. Ez a műszaki-technológiai feltételek szerint az a legáltalánosabb eset, amely a gyakorlatban előfordulhat.

Mindkét esetre példát adunk:

Tegyük fel, hogy valamely erőműben pl. 5 erőművi egység van. Ezek különböző lehetséges együttműködésével adódó üzemmódoknak megfelelő $J(i, j)$ halmazok legyenek a következők:

$$\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 4\}, \{1, 3, 5\}, \{1, 2, 3, 5\}, \{1, 2, 3, 4, 5\}.$$

Az első esetre példa egy olyan tervezési időszak, amelyben ezek közül a következő üzemmódok működése lehetséges:

$$\{1\}, \{1, 2\}, \{1, 2, 3, 5\}, \{1, 2, 3, 4, 5\}.$$

Igaz ugyanis, hogy:

$$\{1, 2, 3, 4, 5\} \supset \{1, 2, 3, 5\} \supset \{1, 2\} \supset \{1\}.$$

A második esetre példa egy olyan tervezési időszak, amelyben a következő üzemmódok lehetségesek:

$$\{1\}, \{1, 2\}, \{1, 3\}, \{1, 3, 5\}, \{1, 2, 3, 5\} \text{ és } \{1, 2, 3, 4, 5\}.$$

Ekkor ugyanis

$$\{1, 2, 3, 4, 5\} \supset \{1, 2, 3, 5\} \supset \{1, 2\} \supset \{1\}; \quad \{1, 3, 5\} \supset \{1, 3\}$$

és

$$\{1, 2, 3, 4, 5\} \supset \{1, 3, 5\}; \quad \{1, 3\} \supset \{1\}$$

teljesül.

Olyan tervezési időszak, amelyben az utóbbi példában szereplő üzemmódokon túl pl. még az $\{1, 2, 4\}$ erőművi egység indexhalmazzal jellemzett üzemmód is megen-

gedett, műszaki okok miatt nem fordulhat elő. Lehetséges azonban egy olyan tervezési időszak, amikor az $\{1\}$, $\{1, 2\}$, $\{1, 2, 4\}$ üzemmódok megengedettek. Ez ismét az első esetre szolgálna példával.

Mivel ez a cikk egy adott ütemezési időszakra vonatkozó modellel foglalkozik, a leírásban szereplő $M(i)$, $i=1, 2, \dots, K$, az adott ütemezési időszakban alkalmazható üzemmódok számát jelöli és nem az összes lehetséges üzemmódot, így az állásköltség számításakor kihasználható a $J(i, j)$, $j=1, 2, \dots, M(i)$, $i=1, 2, \dots, K$ halmazok fenti tulajdonsága.

Az egyszerűsített modellben feltételezzük, hogy az üzemmódok megadásának a sorrendje olyan, hogy: $J(i, 1) \supset J(i, 2) \supset \dots \supset J(i, M(i))$.

3.2.3. Hálózati veszteségből eredő költségek

Az erőművekben termelt villamos energia a hálózaton keresztül jut el a fogyasztókhoz. A szállítás a hálózat ágainak ohmos ellenállása következtében hatásosteljesítmény-veszteséggel jár, a villamos energia egy része hővé alakul. Egy rögzített periódust véve a hálózati veszteségből eredő költség a következő alakú:

$$(3.10) \quad C^v(v, w) = \gamma a_t P^v(v, w),$$

ahol $C^v(v, w)$ a hálózati energia-veszteség forintértéke az adott periódusra, γ egy MWh költsége forintban, a_t a t -edik periódus hossza órában, $P^v(v, w)$ pedig az összes hálózati hatásosteljesítmény-veszteség MW-ban. A $P^v(v, w)$ függvény a v, w változók konvex kvadratikusan függvénye (lásd (F4. 21), (F4. 30), (F4. 31)). A teljes hálózati veszteség-költség az egyes periódusok veszteség-költségeinek összege.

3.3. Korlátozó feltételrendszer

A modell feltételrendszere egyrészt olyan feltételeket tartalmaz, amelyek megkövetelése a változók definíciója miatt szükséges, másrészt olyanokat, amelyek az energiarendszer működésének leírásához szükségesek. A feltételek csoportosíthatók aszerint, hogy periódusonként ismétlődően csak egy-egy periódushoz tartozó változók közötti kapcsolatot fejeznek ki, vagy több periódus változóinak közötti összefüggést írnak elő.

3.3.1. Periódusonként ismétlődő feltételek

Minden egyes periódusban meg kell követelnünk, hogy minden erőműben egyetlen üzemmód működjön, az üzemmód teljesítményszintje a megengedett tartományba essen és a termelt villamos energia mennyisége az országos villamosenergia-igénynek a hálózati veszteséggel és a villamos energia termeléséhez szükséges önfogyasztással megnövelt értékével egyenlő legyen. Meg kell követelnünk azt is, hogy a villamosenergia-rendszer hálózati viszonyai minden periódusban megengedettek legyenek.

Az üzemmód változók definíciója miatt a „minden erőműben egyetlen üzemmód” követelmény a (3.1) egyenlőségek teljesülését jelenti. Ezekre a feltételekre a diszkrét programozásban elfogadott terminológiát fogjuk alkalmazni, *SOS feltételekként*

hivatkozunk rájuk. (*Special Ordered Set* olyan 0 vagy 1 értékű változók halmaza, melyek közül egy és csak egy lehet 1 értékű.)

Természetesen az üzemmód változókra, definíciójuk miatt ezen kívül még az

$$y'_{ij} \in \{0,1\}, \quad j = 1, 2, \dots, M(i), \quad i = 1, 2, \dots, K$$

követelményeknek teljesülniük kell.

A teljesítmény változók definíciója szerint minden periódusban meg kell követelnünk a (3.2) egyenlőtlenségek teljesülését. Ezekre az egyenlőtlenségekre a továbbiakban *összekapcsolási feltétel* elnevezéssel hivatkozunk. Ezek ugyanis azok a feltételek, amelyek közvetlenül „összekapcsolják” az egy-egy erőmű egy-egy üzemmódjához tartozó üzemmód és teljesítmény változót.

Az „energiaigény kielégítés” követelményét tartalmazó feltétel — (minden periódushoz egyetlen ilyen szükséges) — neve *ellátási feltétel*. A teljesítmény és üzemmód-változók felhasználásával a villamosenergia-rendszer t -edik periódusbeli termelésének a szintje a

$$(3.11) \quad \sum_{i=1}^K \sum_{j=1}^{M(i)} (P_{ij}^{\min} y'_{ij} + p'_{ij})$$

összeggel adható meg. Minden erőmű minden üzemmódjához ismert egy ún. önfogyasztás függvény, amely az üzemmód teljesítményszintjének függvényében megadja az erőmű üzemeltetéséhez szükséges villamos energia mennyiségét. Jelölje $P_{ij}^{\text{önf}}$ az i -edik erőmű j -edik üzemmódjához tartozó önfogyasztás függvényt. Az i -edik erőmű önfogyasztása a t -edik periódusban:

$$(3.12) \quad \sum_{j=1}^{M(i)} P_{ij}^{\text{önf}} (P_{ij}^{\min} y'_{ij} + p'_{ij}),$$

és a teljes villamoserőmű-rendszer t -edik periódusbeli önfogyasztása

$$(3.13) \quad P^{t\text{önf}} = \sum_{i=1}^K \sum_{j=1}^{M(i)} P_{ij}^{\text{önf}} (P_{ij}^{\min} y'_{ij} + p'_{ij}).$$

Jelölje $P^{t\text{igény}}$ a t -edik periódusbeli villamosenergia-igény értékét és $P^{t\text{veszt}}$ a t -edik periódusbeli hálózati veszteséget.

Az *ellátási feltétel* a következő:

$$(3.14) \quad \sum_{i=1}^K \sum_{j=1}^{M(i)} (P_{ij}^{\min} y'_{ij} + p'_{ij}) = P^{t\text{igény}} + P^{t\text{veszt}} + P^{t\text{önf}}.$$

Megjegyezzük, hogy a $P^{t\text{veszt}}$ és a $P^{t\text{önf}}$ értékek nemlineárisan függenek a p' és y' változóktól, míg $P^{t\text{igény}}$ konstans érték, az igénygörbe adata. A veszteség és az önfogyasztás számításának módját itt nem részletezzük. A 4. fejezetben egyszerűsítő feltevéseket teszünk ezekre vonatkozóan is.

A villamosenergia-rendszer hálózati viszonyait leíró feltételek a következők (ezek leírásában a periódust azonosító felső indexet elhagyjuk).

Az átviteli hálózattal kapcsolatos feltételrendszer felépítését a csomóponti teljesítmény-betáplálásokat a csomóponti feszültségekkel összekötő (F3. 43) összefüggésekből kiindulva kezdjük el, amelyeket itt a tárgyalás könnyítése céljából reprodukálunk:

$$(3.15) \quad \left. \begin{aligned} f_i(\mathbf{v}, \mathbf{w}) &= P_i \\ g_i(\mathbf{v}, \mathbf{w}) &= Q_i \end{aligned} \right\} \quad i = 1, \dots, N.$$

Jelölje a továbbiakban I_E azon csomópontok sorszámainak halmazát, amelyekhez erőmű csatlakozik, I_E elemszámát jelöljük N_E -vel. Jelöljük továbbá I_M -mel azon csomópontok sorszámainak halmazát, amelyekhez a villamosenergia-rendszer olyan berendezése csatlakozik, amelynek a meddőteljesítmény-betáplálása, illetve fogyasztása az adott periódusban nem előírt értékű, megadott korlátok között mozoghat. Ezeket a továbbiakban meddőforrás-csomópontoknak fogjuk nevezni. (Részletesebben lásd F3. rész). Az I_M halmaz elemszámát jelölje N_M . Az I_E és I_M indexhalmazokat az elemeket alkotó sorszámszámok növekvő sorrendje szerint rendezettnek tekintjük.

A (3.15) egyenlőségekben szereplő P_i hatásos teljesítményt a következőképpen felbontva kezeljük:

$$(3.16) \quad P_i = P_i^G - P_i^F, \quad i = 1, \dots, N,$$

ahol a G felső index generált (betáplált) teljesítményt, az F pedig fogyasztást jelent. A P_i^F , $i=1, \dots, N$ mennyiségek a csomópontként megadott fogyasztás-értékek (teljesítményigények), és nyilvánvalóan teljesül a $P_i^G=0$, $i \notin I_E$ reláció.

Jelölje $\hat{\mathbf{P}}^G$ a \mathbf{P}^G vektor $i \in I_E$ komponenseiből alkotott N_E dimenziós vektort. A (3.15) egyenlőségeknek a modellbe feltételként való beépítéséhez elegendő a $\hat{\mathbf{P}}^G$ vektort a modellnek az erre a periódusra vonatkozó teljesítmény, ill. üzemmód változóiból alkotott \mathbf{p} és \mathbf{y} vektorokkal kifejezni. Jelöljük n -nel \mathbf{p} (és ezzel együtt \mathbf{y}) dimenzióját.

Az üzemmódonként ténylegesen termelt teljesítmény (jelölje $\hat{\mathbf{p}} \in R^n$) ekkor a következőképpen kapható meg:

$$(3.17) \quad \hat{\mathbf{p}} = \mathbf{p} + \text{diag}(\mathbf{p}^{\min})\mathbf{y}.$$

ahol $\text{diag}(\mathbf{p}^{\min})$ olyan diagonális mátrix, melynek diagonális elemei a \mathbf{p}^{\min} komponensei, \mathbf{p}^{\min} erőművenként és üzemmódonként a minimális teljesítményszintek vektora.

Az egyes csomópontokhoz tartozó hatásosteljesítmény-betáplálás ezek után egyszerűen úgy kapható meg, hogy az illető csomópont-hoz csatlakozó erőművek üzemmódjaira a ténylegesen termelt teljesítményeket összeadjuk, azaz $\hat{\mathbf{p}}$ megfelelő indexű komponenseit összeadjuk. Mindez a következő lineáris transzformáció végrehajtását jelenti:

$$(3.18) \quad \hat{\mathbf{P}}^G = \mathbf{H}\hat{\mathbf{p}}.$$

Itt \mathbf{H} ($N_E \times n$) méretű összegező mátrix, i -edik sorában az i -edik erőművi csomópont-hoz csatlakozó erőmű üzemmódjainak megfelelő pozíciókban 1 áll, a sorban a többi elem 0; $i=1, \dots, N_E$.

Jelöljük a (3.17), (3.18) által definiált lineáris transzformáció mátrixát \mathbf{D} -vel, ennek mérete $N_E \times (2n)$ lesz. Legyen továbbá $\mathbf{D}^T = (\mathbf{d}_1, \dots, \mathbf{d}_L)$. Így adódik az aláb-

bi összefüggés:

$$(3.19) \quad \hat{\mathbf{p}}^G = \mathbf{D} \begin{pmatrix} \mathbf{p} \\ \mathbf{y} \end{pmatrix}.$$

Ha $i \in I_E$, jelölje $E(i)$ az i index sorszámát az I_E sorbarendezett halmazon belül.

Mielőtt a (3.15) feltételrendszernek a modellhez csatolandó végleges alakját megadnánk, még egy megjegyzést teszünk. A (3.15) egyenlőségekben a Q_i mennyiségek $i \in I_M$ esetén változóként kezelendők. (Lásd Függelék, F3. rész.) Mivel azonban a reaktív teljesítményre vonatkozó feltételekben ezek explicitálva vannak, ezért az $i \in I_M$ sorszámú feltételek elhagyhatók a reaktív teljesítményre vonatkozó részből.

Ezek után a \mathbf{p} , \mathbf{y} változókat a csomóponti feszültségekkel összekötő feltételrendszer a következő alakot ölti:

$$(3.20) \quad \begin{aligned} \mathbf{d}_{E(i)}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{y} \end{pmatrix} - f_i(\mathbf{v}, \mathbf{w}) &= P_i^F, \quad i \in I_E \\ -f_i(\mathbf{v}, \mathbf{w}) &= P_i^F, \quad i \notin I_E \\ g_i(\mathbf{v}, \mathbf{w}) &= Q_i, \quad i \in I_M, \end{aligned}$$

ahol az $f_i(\mathbf{v}, \mathbf{w})$, $g_i(\mathbf{v}, \mathbf{w})$ függvények a \mathbf{v}, \mathbf{w} változó kvadratikus függvényei (lásd (F3. 41)), a feltételek jobb oldalán konstans mennyiségek szerepelnek.

A hálózattal kapcsolatos feltételrendszer további részei a különböző villamos mennyiségekre vonatkozó egyenlőtlenség alakú korlátozásokat fejeznek ki, ezeket egyenként áttekintjük.

Feszültségtartási feltételek

A villamosenergia-rendszerben valamennyi fogyasztó rögzített névleges feszültségszintre van tervezve, az ettől való túl nagy eltérés a fogyasztó tönkremenésével, vagy hibás működésével jár. A fogyasztók feszültségének a névleges szint közelében való tartásához szükség van arra, hogy a nagyfeszültségű hálózat csomópontjainak feszültsége adott intervallumba essen. Legyen a nagyfeszültségű átviteli hálózatra a feszültség abszolút értékének alsó korlátja az i -edik csomópontra vonatkozóan V_i^{\min} , a felső korlát V_i^{\max} , $i = 1, \dots, N$.

Ekkor a feszültségtartási feltételek a következők:

$$(3.21) \quad V_i^{\min} \leq (v_i^2 + w_i^2)^{1/2} \leq V_i^{\max}, \quad i = 1, \dots, N.$$

Ágterhelési feltételek

Ezek az ágak termikus terhelhetőségére vonatkoznak, a túlmelegedést korlátozzák. A hálózat minden (i, k) ágához adott egy T_{ik}^{\max} érték, az ágon áramló teljesítmény abszolút értéke (látszólagos teljesítmény) biztonsági okokból ennél nem lehet nagyobb. A teljesítményáramlást a Függelék 4. részében tárgyaljuk, a látszólagos teljesítményt a modell változóival az (F4. 20) egyenlőség köti össze. Az ágon létrejövő veszteséget

ebben a vonatkozásban elhanyagolhatjuk, így az ágterhelési feltételek az alábbiak lesznek:

$$(3.22) \quad |S_{ik}(v_i, v_k, w_i, w_k)| \leq T_{ik}^{\max},$$

az átviteli hálózat minden (i, k) ágára.

Meddőforrás feltételek

E feltételek a meddőforrás-csomópontok (esetünkben az $i \in I_M$ sorszámuak) meddőteljesítmény-kibocsátását, ill. nyelését korlátozzák. Erőműveket, ill. szinkron-kompenzátorokat tekintve arról van szó, hogy a szinkron-generátorok karakterisztikáitól függően azok meddőteljesítmény-kibocsátása (ill. nyelése) korlátozott, amiből korlátok adódnak a megfelelő csomópont meddőteljesítmény-kibocsátására, ill. nyelésére vonatkozóan. A többi meddőforrásra vonatkozóan is alsó és felső korlátok adódnak a meddőteljesítmény-kibocsátásra (nyelésre). A meddőforrásokkal kapcsolatban lásd F2. részt.

Ha az i -edik csomópont meddőforrás, és a meddő teljesítmény korlátai az erőművi esetben $Q_i^{\min}(\mathbf{y})$, $Q_i^{\max}(\mathbf{y})$, $i \in I_M \cap I_E$, egyébként Q_i^{\min} , Q_i^{\max} , $i \in I_M \setminus I_E$, akkor a következő feltételek adódnak:

$$(3.23) \quad \begin{aligned} Q_i^{\min}(\mathbf{y}) &\leq Q_i \leq Q_i^{\max}(\mathbf{y}), & i \in I_M \cap I_E \\ Q_i^{\min} &\leq Q_i \leq Q_i^{\max}, & i \in I_M \setminus I_E \end{aligned}$$

Mint azt már említettük, a (3.15) egyenlőségekben a Q_i mennyiségek explicitált alakban szerepelnek, így behelyettesítés után adódnak a meddőforrás feltételek:

$$(3.24) \quad \begin{aligned} Q_i^{\min}(\mathbf{y}) &\leq g_i(\mathbf{v}, \mathbf{w}) \leq Q_i^{\max}(\mathbf{y}), & i \in I_M \cap I_E \\ Q_i^{\min} &\leq g_i(\mathbf{v}, \mathbf{w}) \leq Q_i^{\max}, & i \in I_M \setminus I_E. \end{aligned}$$

A (3.24) egyenlőtlenségekben az erőműveknek megfelelő sorokban a jobb és bal oldalon szereplő korlátok az üzemiállapot változók függvényei, ugyanis a csomóponti meddőteljesítmény-kibocsátásra vonatkozó korlátok a csomópontokhoz csatlakozó erőművek üzemmódjainak meddő-kibocsátására vonatkozó korlátokból összeadás révén keletkeznek. Jelölje \hat{Q}_i^{\min} , ill. \hat{Q}_i^{\max} a következő vektorokat: dimenziójuk megegyezik \mathbf{y} dimenziójával, az i -edik csomópontokhoz csatlakozó erőművek üzemmódjainak pozíciójában a megfelelő üzemmódra vonatkozó alsó, ill. felső korlátok vannak, a többi pozícióban pedig 0. Így a fent említett összeadás megvalósítása és ezzel a meddőforrás korlátok a következők:

$$Q_i^{\min}(\mathbf{y}) = \mathbf{y}^T \hat{Q}_i^{\min}, \quad Q_i^{\max}(\mathbf{y}) = \mathbf{y}^T \hat{Q}_i^{\max}, \quad i \in I_M \cap I_E.$$

3.3.2. Periódusokat összekapcsoló feltételek

A modell több különböző periódusának változói között az ún. *kikapcsolási feltételek* és a *tüzelőanyagkényszer feltételek* létesítenek kapcsolatot.

Jellemző, hogy a kikapcsolási feltételekben csak az üzemmód változók szerepelnek, és a tüzelőanyagkényszer feltételek mindegyike egyetlen erőmű teljesítmény és

üzemmód változóit tartalmazza. Olyan feltételek, amelyek több periódus feszültségi változóit tartalmazzák, nincsenek a modellben.

A *kikapcsolási feltételekben* azt a követelményt fogalmazzuk meg, hogy műszaki okok következtében a kikapcsolt egység a minimális, 4 órás állásidő eltelte előtt nem térhet vissza a termelésbe, gazdasági okokból pedig az erőműnek a kikapcsolttal azonos egysége sem indíthat. Ez utóbbi ugyan műszakilag lehetséges lenne, de gazdasági szempontból nem indokolt egy egység kikapcsolása, ha például 2 óra múlva vele azonos jellegű és teljesítményű egység indítása válik szükségessé.

A kikapcsolt egységek minimálisan 4 órás állásidejét biztosító követelmény a 3.2.2. bekezdésben definiált $\xi_{ik}(t)$ állásidő segítségével úgy fogalmazható meg, hogy a

$$(3.25) \quad \xi_{ik}(t-1) \left(1 - \sum_{j:k \in J(i,j)} y_{ij}^{t-1}\right) \left(\sum_{j:k \in J(i,j)} y_{ij}^t\right)$$

szorzat értéke vagy 0 vagy legalább 4. Ennek $t=1, 2, \dots, T$, $i=1, 2, \dots, K$ és $k=1, 2, \dots, N(i)$ -re teljesülnie kell. A feltétel ilyen megfogalmazása előnyös, mert ebben nem kell külön kezelni az előző tervezési időszak végén, ill. végét megelőzően leállított egységeket.

Ugyanakkor azonban $\xi_{ik}(t)$ számítása bonyolult, s ezért célszerű ezt a követelményt ezen mennyiségek felhasználása nélkül megfogalmazni.

Vezessük be a következő jelölést. Legyen rögzített t esetén ($1 \leq t \leq T-1$) $l(t)$ az az index, amelyre

$$t + l(t) \leq T,$$

$$a_t + a_{t+1} + \dots + a_{t+l(t)} \geq 4,$$

$$a_t + a_{t+1} + \dots + a_{t+l(t)-1} < 4.$$

Ha ilyen $l(t)$ nem található, akkor legyen $l(t) = T - t$. E szerint a definíció szerint $l(t)$ a t -edik periódussal kezdődően azon periódusok száma, amelyek együttes időtartama már eléri a 4 órát, feltéve, hogy a t -edik periódust követő 4 óra (beleértve a t -edik periódus a_t időtartamát is) még a tervezési időszakhoz tartozik. Ha ez nem teljesül, akkor $l(t)$ a tervezési időszak t -edik periódus utáni periódusainak száma. E jelölés bevezetésére azért van szükség, mert a vizsgált időszakban vannak nem egy-órás periódusok is, és a követelmény úgy szól, hogy a kikapcsolt egységeknek legalább 4 órán át kell állnia.

A „minimálisan 4 órás állásidő” követelmény így azt írja elő, hogy minden $i=1, 2, \dots, K$ és $k=1, 2, \dots, N(i)$, valamint $t=1, 2, \dots, T-1$ értékre

$$\sum_{j:k \in J(i,j)} y_{ij}^{t-1} = 1 \quad \text{és} \quad \sum_{j:k \in J(i,j)} y_{ij}^t = 0$$

teljesülésekor

$$\sum_{j:k \in J(i,j)} y_{ij}^{t+l} = 0$$

legyen, $l = 1, 2, \dots, l(t)$ esetén.

Könnyen ellenőrizhető, hogy ennek biztosítására alkalmas a következő egyenlőtlenség-rendszer:

$$(3.26) \quad \sum_{j: k \in J(i, j)} y_{ij}^{t-1} - \sum_{j: k \in J(i, j)} y_{ij}^t + \sum_{j: k \in J(i, j)} y_{ij}^{t+l} \leq 1, \\ t = 1, 2, \dots, T-1, \\ l = 1, 2, \dots, l(t), \\ i = 1, 2, \dots, K, \\ k = 1, 2, \dots, N(i).$$

Ezek a feltételek azt írják elő, hogy ha valamely erőművi egység kikapcsolásra került a tervezési időszak egy periódusában, akkor ne kerüljön bekapcsolásra az azt követő 4 órán belül. Fel kell írunk még azokat a feltételeket is, amelyek azt biztosítják, hogy a tervezési időszakot megelőző 4 órán belül kikapcsolt erőművi egység bekapcsolása is csak a szükséges 4 órás állásidő után következhesen be. A kikapcsolási feltételek közül ezeket *első-indítási feltételeknek nevezzük*.

Megfogalmazásukhoz legyen l_{ik} értéke a tervidőszak kezdetén az i -edik erőmű k erőművi egységénél szükséges továbbállási periódusok száma, azaz $\sum_{j: k \in J(i, j)} y_{ij}^0 = 0$ és $\xi_{ik}(0) < 4$ esetén l_{ik} annak a periódusnak a sorszáma, amelyre

$$\xi_{ik}(0) + a_1 + \dots + a_{l_{ik}} \cong 4$$

és

$$\xi_{ik}(0) + a_1 + \dots + a_{l_{ik}-1} < 4,$$

egyébként pedig $l_{ik} = 0$.

Minden erőmű minden egységére a tervezési időszak első l_{ik} periódusában meg kell követelnünk, hogy

$$\sum_{j: k \in J(i, j)} y_{ij}^t = 0$$

legyen. Így az első-indítási feltételek a következők:

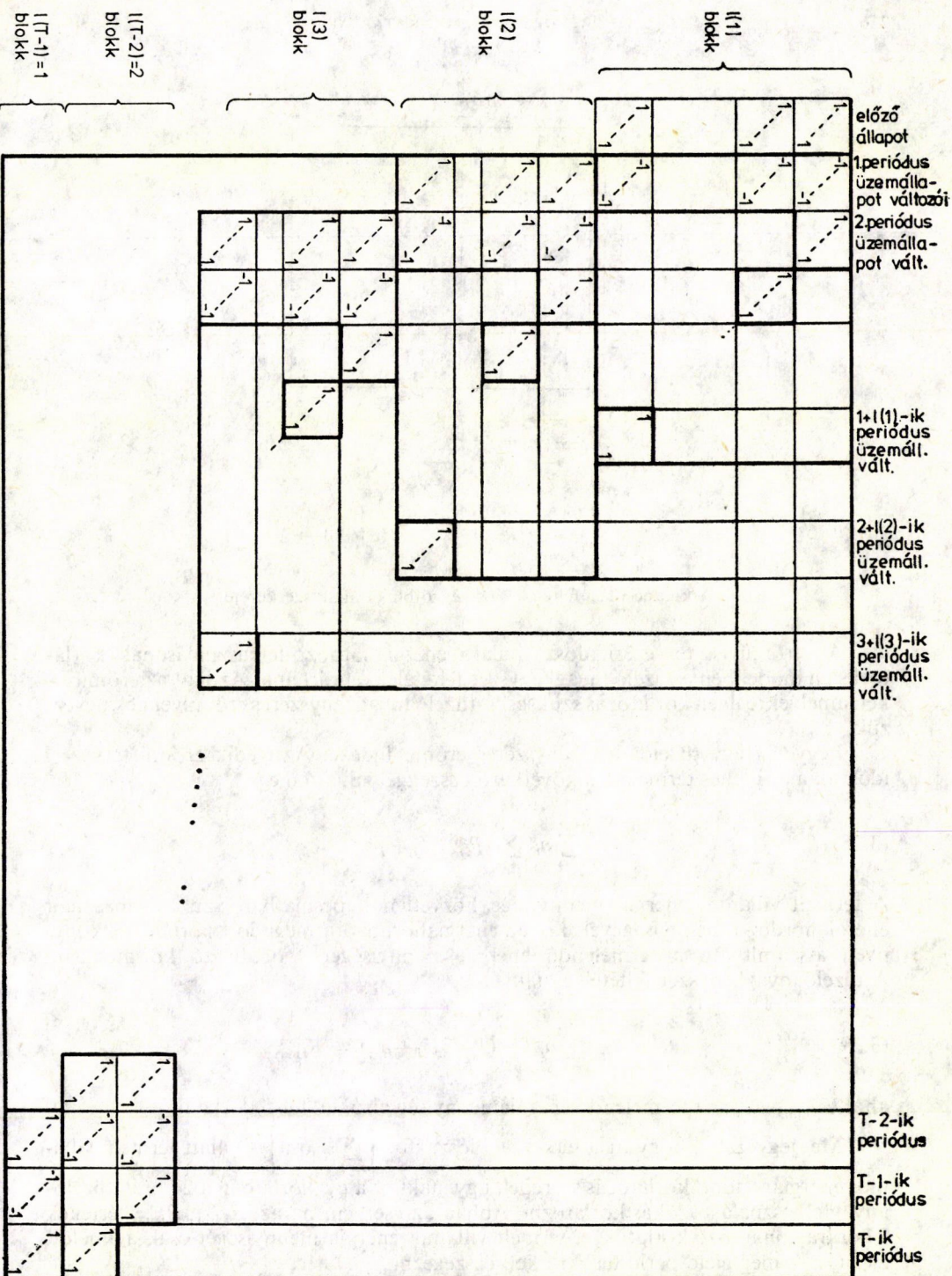
$$(3.27) \quad \sum_{j: k \in J(i, j)} y_{ij}^t = 0, \quad i = 1, 2, \dots, K, \quad l = 1, 2, \dots, l_{ik}, \quad k = 1, 2, \dots, N(i).$$

Emlékeztetjük az olvasót arra, hogy a 3.1.1. bekezdésben szereplő megjegyzés szerint az üzemmód változókra egyéb, egyedi korlátozás is szükséges lehet. Ezen kívül előfordulhat még az is, hogy bizonyos okok miatt egy-egy erőműben bizonyos üzemmódok nem alkalmazhatóak közvetlenül egymás után. E kérdéssel azonban itt most nem foglalkozunk.

A 6. ábrán a kikapcsolási feltételek szerkezetét tekintjük át, mátrixos sémával. (Az első-indítási feltételek ábrázolásától eltekintünk.) Az ábrán egység-blokkal jelölt, valójában összegző blokkok szerkezetét a 7. ábrán külön is bemutatjuk.

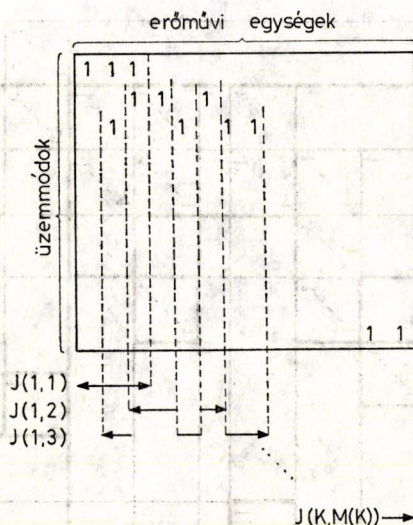
Több periódus változói közötti kapcsolatot írnak elő a *tüzelőanyagkényszer feltételek* is.

A 2.2.2. bekezdésben leírtak szerint a villamosenergia-rendszer zavartalan működéséhez szükséges lehet, hogy valamely erőmű vagy erőművek energiahordozó-felhasználását korlátozzuk. A korlátozás lehet az energiahordozó mennyiségére vonatkozó alsó korlát vagy felső korlát, esetleg alsó és felső korlát vagy egyenlőség követelmény.



6. ábra

A kikapcsolási feltételek áttekintése mátrixos sémával



7. ábra

A kikapcsolási feltételek összegző-blokkjának szerkezete

Az erőművek tervezési időszak alatti energiahordozó-felhasználásának korlátozása a modellben a tüzelőanyagkényszer feltételekkel történik. Azokat az erőműveket, amelyekre ilyen korlátozás szükséges, tüzelőanyagkényszeres erőműveknek nevezzük.

Legyen i egy tüzelőanyagkényszeres erőmű indexe. Az i -edik erőmű tervezési időszak alatti teljes termelése a következő összeggel adható meg:

$$(3.28) \quad \sum_{t=1}^T a_t \sum_{j=1}^{M(i)} (P_{ij}^{\min} y_{ij}^t + p_{ij}^t).$$

A termelt villamos energia mennyisége közvetlen kapcsolatban van a felhasznált energiahordozó mennyiségével. Így az energiahordozóra megadott korlátozás könnyen átszámítható a termelendő energia mennyiségére vonatkozó korlátozásra. A tüzelőanyagkényszer feltételek alakja:

$$(3.29) \quad R_{i\min} \leq \sum_{t=1}^T a_t \sum_{j=1}^{M(i)} (P_{ij}^{\min} y_{ij}^t + p_{ij}^t) \leq R_{i\max},$$

ahol $R_{i\min}$, $R_{i\max}$ a termelendő energiamennyiség alsó, ill. felső korlátai.

Megjegyezzük, hogy itt a tervezési időszak — $\left(\sum_{t=1}^T a_t \text{ óra}\right)$ — alatt termelt villamos energiára adott korlátozás szerepel, ugyanakkor a gyakorlatban a 24 órás tüzelőanyagfelhasználást szokás korlátozni. Abban az esetben, ha az $R_{i\min}$, $R_{i\max}$ értékek 24 órára vonatkozó korlátok, a termelt villamos energia mennyiségét csak a kijelölt 24 órának megfelelő periódusokra kell összegezni.

3.4. A modell szerkezetének áttekintése, mérete, jellemzői

Az általános modell ismertetésének befejezéseként ebben a bekezdésben röviden áttekintjük a modellnek megfelelő vegyesváltozós, nemlineáris programozási feladat szerkezetét, összefoglaljuk jellemzőit.

A modell változói az y üzemmód, p teljesítmény és a v, w feszültségi változók. Az üzemmód változó komponensek 0 vagy 1 értéket vehetnek csak fel, a teljesítmény változók nem negatívak. A minimalizálandó célfüggvény (3.4) termelési-költség része szeparábilis, nemlineáris függvénye az üzemmód és teljesítmény változó komponenseiből alkotott $P_{ij}^{min} y_{ij}^t + p_{ij}^t$ változóknak. A (3.8) állásköltség az egy-egy erőmű egy-egy üzemmódjához tartozó $y_{ij}^t, t=1, 2, \dots, T$ komponensek nemlineáris függvényeinek az összege. A (3.10) veszteségeköltség a v^t, w^t változók konvex kvadrátikus függvénye, periódusonként.

A korlátozó feltételrendszer strukturált. Minden periódushoz vannak csak az adott periódus változóit tartalmazó feltételek. Így az egyes periódusok üzemmód változójának komponenseire a (3.1) SOS feltételeknek kell teljesülniük, y^t és p^t között teremtenek kapcsolatot a (3.2) összekapcsolási feltételek. A (3.14) ellátási feltételben a

$\sum_{j=1}^{M(i)} (P_{ij}^{min} y_{ij}^t + p_{ij}^t)$ nemlineáris függvénye és a szintén nemlineáris hálózati veszteség szerepel, amely a v^t, w^t változók függvénye.

A (3.21), (3.22), (3.24) hálózati viszonyokat leíró feltételek szintén nemlineárisak, a (3.26), (3.27) kikapcsolási feltételek több periódus üzemmód változóit tartalmazzák, felépítésük a 6. ábrán látható. Egy-egy tüzelőanyagkényszerfeltétel ((3.29)) szintén több periódus változóit tartalmazza, de csak egyetlen erőműhöz tartozó üzemmód és teljesítmény változók lineáris függvénye.

A 8. ábrán a modell vázlatos rajza látható, a változók felsorolásánál egymás után szerepelnek egy adott periódus teljesítmény, üzemmód és feszültségi változói.

Az általános modellnek megfelelő vegyesváltozós, nemlineáris programozási feladat mérete:

Változóinak száma:

1. 0–1 értékű változóinak száma = T -szer az energiatermelő rendszerben alkalmazható üzemmódok száma, azaz

$$T \sum_{i=1}^K M(i);$$

2. teljesítmény változók száma = T -szer az energiatermelő rendszerben alkalmazható üzemmódok száma,

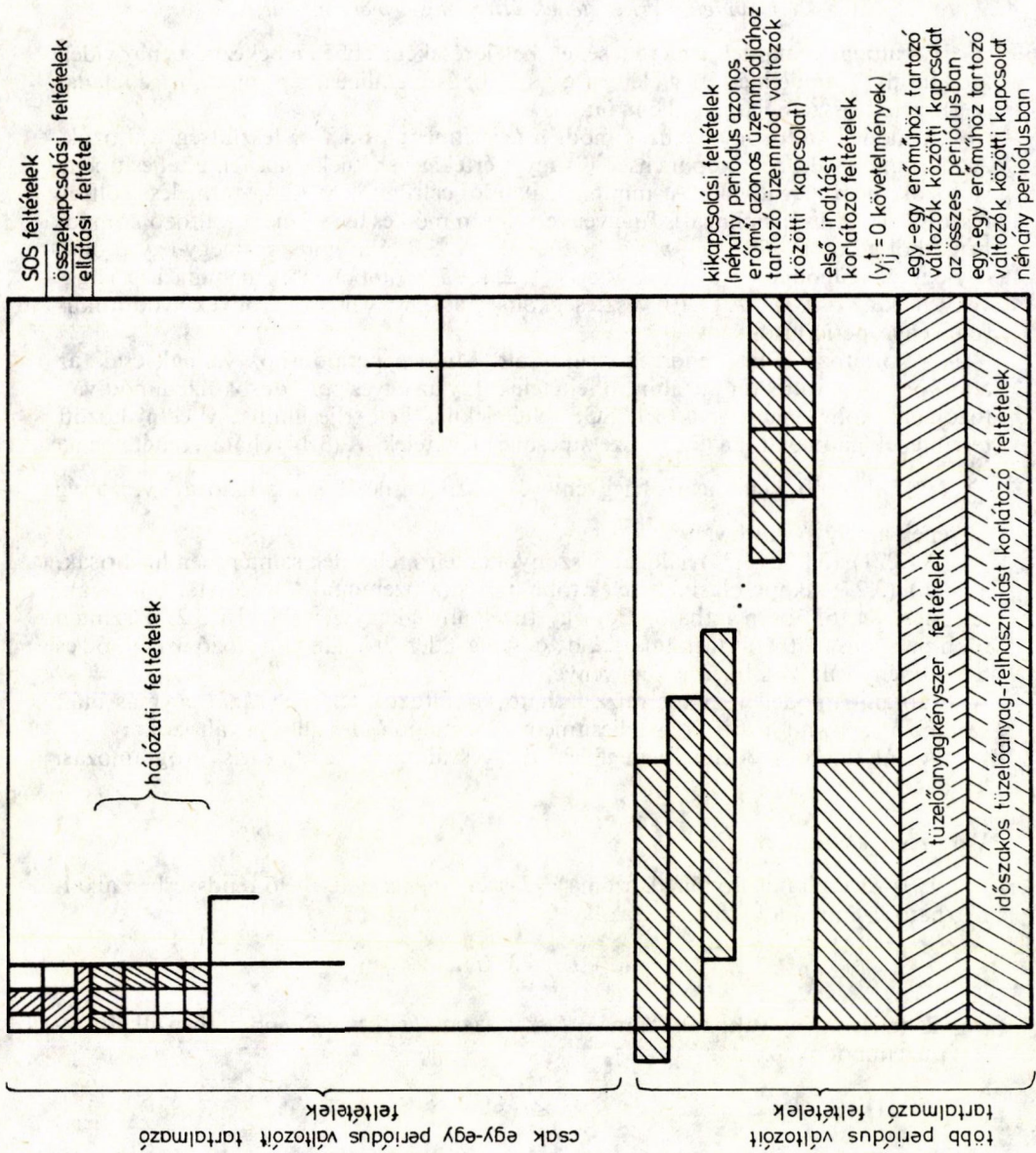
$$T \sum_{i=1}^K M(i);$$

3. folytonos, feszültségi változóinak száma = T -szer a hálózat csomópontjainak a száma,

$$TN.$$

Feltételeinek száma:

1. periódusonként ismétlődően: az erőművek számával megegyező számú SOS feltétel, az energiatermelő rendszerben alkalmazható üzemmódok számával azonos



8. ábra
Az általános
modell
struktúrája

számú összekapcsolási feltétel, 1 ellátási feltétel és a hálózat csomópontjainak kétszerese-számú feszültségtartási, az ágak számának kétszerese-számú ágterhelési feltétel, a meddő csomópontok számával azonos számú meddőforrás feltétel, ezen kívül a hálózatszámítási feladat feltételei.

2. a periódusokat összekapcsoló kikapcsolási feltételek száma az erőműrendszer erőművi egységei számának kb. $4T$ -szerese, ezen kívül van még maximum 5 tüzelőanyagkényszer feltétel.

Megemlítjük még, hogy *ha* e feladat megoldására *leszámlálási jellegű stratégiát alkalmaznánk*, — azaz a feladatot rögzített üzemmód változó értékek mellett oldanánk meg, akkor az így egy-egy lépésben megoldandó feladat mérete a következő lenne:

Változóinak száma: $T(K + N)$

Feltételeinek száma:

1. periódusonként ismétlődően: az erőművek számával azonos számú alsó-felső korlát a teljesítmény változókra (ezek felelnek meg az összekapcsolási feltételeknek), egy ellátási feltétel és a korábbihoz képest változatlan számú hálózati feltétel.

2. a periódusok összekapcsolására ebben a feladatban a tüzelőanyagkényszer feltételek szolgálnának, számuk maximum 5.

Ha a tüzelőanyagkényszer feltétel nem szerepel egy konkrét tervezési időszakra vonatkozóan, akkor ez a feladat szétesik T független feladatra (hiszen az üzemmód változók rögzítettek!). Ezek mindegyike az erőművek számával azonos számú teljesítmény változót és a hálózat csomópontjainak kétszerese számú feszültségi változót tartalmaz. Célfüggvénye a teljesítmény változók konvex, szeparábilis, nemlineáris függvénye. A feltételek között is vannak nemlineárisak, például a hálózati viszonyokat leíró feltételek kvadratikusak, de nem biztos, hogy konvexek.

3.5. A 3. fejezetben alkalmazott jelölések összefoglalása

T	a tervezési időszak periódusainak száma,
a_t	a t -edik periódus időtartama, órában,
K	erőművek száma,
$M(k)$	$k=1, 2, \dots, K$ az egyes erőművekben a tervezési időszakban alkalmazható üzemmódok száma,
i, j	indexpár az i -edik erőmű j -edik üzemmódjára utal, $i=1, 2, \dots, K$, $j=1, 2, \dots, M(k)$,
y^t, y_i^t, y_{ij}^t	a t -edik periódushoz tartozó üzemmódváltozó vektor és komponensei,
y	a modell üzemmódváltozó vektora (y^t vektorok egymás után fűzésével adódik),
y^0	a vizsgált időszakot megelőző legutolsó periódus üzemmódváltozó vektorának értéke,
p^t, p_i^t, p_{ij}^t	a t -edik periódushoz tartozó teljesítmény változó és komponensei,
p	a modell teljesítményváltozó vektora (a p^t vektorok egymás után fűzésével keletkezik),
$p_{ij}^{\min}, p_{ij}^{\max}$	az i -edik erőmű j -edik üzemmódja teljesítményszintjének alsó és felső korlátja,
v^t, w^t	a t -edik periódusban a csomópontok komplex potenciáljának valós, ill. képzetes részéből alkotott vektorok,

$f_{ij}(P)$	az i -edik erőmű j -edik üzemmódjának termelési költségét a teljesítmény függvényében megadó függvény,
$K_{ij} = f_{ij}(P_{ij}^{\min})$	az i -edik erőmű j -edik üzemmódjának működésekor a minimális teljesítmény kibocsátásának költsége,
$k_{ij}(P) = f_{ij}(P_{ij}^{\min} + P) - f_{ij}(P_{ij}^{\min})$	az i -edik erőmű j -edik üzemmódjának működésekor a minimális szint feletti P teljesítmény többlet termelési költsége,
$N(i)$	az i -edik erőmű egységeinek száma,
$J(i, j)$	az i -edik erőmű j -edik üzemmódjának üzemelésekor működő egységek sorszámainak halmaza,
$g_{ik}(\tau)$	az i -edik erőmű k -adik egységének állásköltsége a τ állásidő függvényében,
$G_{ik}(\infty), G_{ik}(0), c_{ik}$	az állásköltség-függvényben szereplő konstansok,
$\tau_{ij}(t)$	az i -edik erőmű j -edik üzemmódja folytatólagos állásának ideje a t -edik periódusig bezárólag,
$\tau_{ij}(0)$	az i -edik erőmű j -edik üzemmódjának a tervezési időszakot megelőző folytatólagos állásideje,
$\xi_{ik}(t)$	az i -edik erőmű k -adik egysége folytatólagos állásának ideje a t -edik periódusig bezárólag,
$\xi_{ik}(0)$	az i -edik erőmű k -adik egységének a tervezési időszakot megelőző folytatólagos állásideje,
$P_{ij}^{\text{önf}}(P)$	az i -edik erőmű j -edik üzemmódja működésekor P szintű termelés esetén fellépő önfogyasztás,
P^{veszt}	a t -edik periódusban fellépő hálózati veszteség (v^t, w^t függvénye),
$P^{\text{önf}}$	az erőműrendszer teljes önfogyasztása a t -edik periódusban (y^t és p^t függvénye),
$P^{\text{t igény}}$	a t -edik periódusban fellépő teljesítményigény, az igénygörbe t -edik periódushoz tartozó értéke (MW),
$l(t)$	a t -edik periódust követő azon periódusok száma, amelyek együttes időtartama — beleértve a t -edik periódus időtartamát is — már eléri a 4 órát, ill. a t -edik periódust követő periódusok száma a tervezési időszakban,
l_{ik}	az i -edik erőmű k -adik egysége csak a tervezési időszak első l_{ik} periódusának eltelte után kapcsolható be először,
$R_{i \min}, R_{i \max}$	tüzelőanyagkényszeres erőmű esetén a tüzelőanyagkényszernek megfelelő termelési korlátok,
N	a hálózat csomópontjainak száma,
M	a hálózat ágainak száma,
I_E	az erőművi csomópontok sorszámaiból alkotott indexhalmaz,
N_E	az I_E elemszáma,
I_M	a meddőforrás csomópontok sorszámainak halmaza,
N_M	az I_M elemszáma,
P_i	a hálózat i -edik csomópontjára a hatásosteljesítmény- betáplálás és fogyasztás különbsége, $i = 1, \dots, N$,
Q_i	a hálózat i -edik csomópontján a meddőteljesítmény- betáplálás és fogyasztás különbsége, $i = 1, \dots, N$,
P_i^G	a hálózat i -edik csomópontján a hatásosteljesítmény- betáplálás, $i = 1, \dots, N$,

P_i^f	a hálózat i -edik csomópontjára a hatásosteljesítmény- fogyasztás, $i=1, \dots, N$,
\hat{P}_i^G	a P^G vektor $i \in I_E$ komponenseiből alkotott N_E -dimenziós vektor,
D	a teljesítmény és üzemmállapot változók, valamint a csomóponti generált teljesítmények közötti lineáris transzformáció mátrixa,
p^{\min}	komponensei erőművenként és üzemmódonként a minimális teljesítményszintek,
v_i	az i -edik csomópont feszültségének valós része,
w_i	az i -edik csomópont feszültségének képzetes része,
V_i^{\min}	az i -edik csomópont feszültsége abszolút értékének alsó korlátja,
V_i^{\max}	az i -edik csomópont feszültsége abszolút értékének felső korlátja,
$Q_i^{\min}(y)$	a meddő teljesítmény alsó korlátja, $i \in I_M \cap I_E$,
$Q_i^{\max}(y)$	a meddő teljesítmény felső korlátja, $i \in I_M \cap I_E$,
Q_i^{\min}	a meddő teljesítmény alsó korlátja, $i \in I_M \setminus I_E$,
Q_i^{\max}	a meddő teljesítmény felső korlátja, $i \in I_M \setminus I_E$,
T_{ik}^{\max}	a látszólagos teljesítmény felső korlátja az (i, k) ágra vonatkozóan (termikus terhelhetőség),
$f_i(v, w)$	az i -edik csomópontához tartozó hatásos teljesítménynek a hálózati feszültségeloszlástól való függését leíró függvény,
$g_i(v, w)$	az i -edik csomópontához tartozó reaktív teljesítménynek a hálózati feszültségeloszlástól való függését leíró függvény,
$S_{ik}(v_i, v_k, w_i, w_k)$	a hálózat i -edik csomópontjáról az (i, k) ágra áramló teljesítmény,
$P^v(v, w)$	a hatásosteljesítmény-veszteség, mint a feszültségeloszlás függvénye.

4. A modellt egyszerűsítő feltevések

Az ütemezési feladat 3. fejezetben ismertetett általános modelljének megfogalmazásakor elsősorban az volt a célunk, hogy egy — a problémát a lehető legjobban leíró — modellt hozzunk létre, függetlenül attól, hogy az ennek megfelelő matematikai programozási feladat megoldására van-e esélyünk. A modell 3.4. bekezdésben szereplő áttekintése alapján a megfelelő feladat nagyméretű, vegyesváltozós, célfüggvényében és feltételeiben is nemlinearitást tartalmazó feladat. Megoldására nincs mód.

Szükségesnek látszik, hogy megkíséreljük a modellben szereplő változók és feltételek számának csökkentését, a nemlinearitások kiküszöbölését úgy, hogy az így létrejövő modell az ütemezési problémának még jó modellje, a megfelelő feladat pedig már kezelhető feladat legyen.

Ebben a fejezetben ismertetjük azokat az egyszerűsítő feltevéseket, amelyek figyelembevételével az általános modell egyszerűsíthető a fenti értelemben.

Az ütemezési probléma modellezésekor ezeknek a feltevéseknek a figyelembevétele *célszerű*, hiszen az így létrejövő modellnek megfelelő matematikai programozási feladat megoldása lehetséges.

Feltételezésük ugyanakkor *megalapozott*, hiszen a gyakorlatban már bevált üzemviteli technológia jellegzetességeit vagy matematikai jellegű közelítéseket tartalmaznak. Használunk pusztán jelöléstechnikai jellegű egyszerűsítő feltevéseket is.

Az 5. fejezet az egyszerűsített modell leírását tartalmazza, a figyelembe vett tervezési időszak 25 órás időtartamú, $T=27$ periódust tartalmaz, $a_t=1$ vagy $0,5$, $t=1, 2, \dots, 27$.

4.1. Az igénygörbe jellegzetes alakján alapuló egyszerűsítő feltevés

Az országos villamosenergia-igény napi alakulását megadó igénygörbe jellegzetes tulajdonsága, hogy két-két maximális igényérték között egymást követően néhány olyan periódus található, amelyben az igény értéke monoton csökken, majd néhány periódus, amelyben egy minimális érték körül kismértékben ingadozik, és végül olyan periódusok következnek, amelyekben az igény monoton nő.

Ésszerűnek tűnik az a követelmény, hogy olyan periódusokban, amelyekben az igény értéke monoton csökken, ne engedjünk meg az erőművekben olyan üzemmódváltást, amely az erőművi egység(ek) bekapcsolását, azaz felfűtését igényli, hanem csak olyat, amely erőművi egység(ek) kikapcsolása által valósul meg.

Olyan periódusokban, amelyekben az igény monoton növekszik, a lehetséges üzemmódváltást korlátozzuk úgy, hogy csak a már működő üzemmódból bekapcsolással megvalósítható üzemmódok legyenek megengedettek.

Azokban a periódusokban, amelyekben az igény csak kismértékben ingadozik, nem engedjük meg az üzemmódváltoztatást, hanem csak az alkalmazott üzemmód teljesítményszintjének a megváltoztatását.

Egyszerűsítő feltételként az előzőek szerint az egynapos tervezési időszakot felbonthatjuk hat egymást követő ún. szakaszra (egymást követő periódusok együttes elnevezése).

Az első és a negyedik szakaszt az jellemzi, hogy az igény monoton csökken. A harmadik és hatodik szakaszt az, hogy a hozzá tartozó periódusokban az igény monoton nő. A második és ötödik szakaszhoz tartozzon 4-4 periódus, amelyekben az igényérték a minimum körül ingadozik.

Nevezzük az első és negyedik szakaszt *kikapcsolási szakasznak*, a harmadik és hatodik szakaszt *bekapcsolási szakasznak*, a második és negyedik szakaszt *stagnálási szakasznak*.

Az optimumkeresés során szorítkozzunk azon menetrendek vizsgálatára, amelyek a kikapcsolási szakaszokban csak kikapcsolás által megvalósítható üzemmódváltásokat, a bekapcsolási szakaszokban csak bekapcsolás által megvalósítható üzemmódváltozásokat írnak elő, a stagnálási szakaszokban pedig nincs üzemmódváltás.

A lehetséges menetrendek halmazának ilyen korlátozása összhangban van az üzemviteli gyakorlattal, s logikusan következik az üzemmódváltás költségkihatásának ismeretéből. Ugyanakkor a modell szempontjából ez egyszerűsítő feltevés, hiszen a feltevésnek eleget tevő menetrendek automatikusan biztosítják a kikapcsolt berendezések minimálisan 4 órás állását. Így a 3.3.2. bekezdésben ismertetett kikapcsolási feltételeket elhagyhatjuk a modellből, a feltételek száma csökkenthető.

Megjegyezzük, hogy más országokban a nálunk alkalmazott üzemviteli szabályokhoz nagy vonalakban hasonló, de apróbb részletekben a mieinktől eltérő megállapodások érvényesülnek [45].

4.2. Az üzemmódok megadásának sorrendje

A 3.2.1. bekezdés végén található megjegyzés szerint az erőművekben egy adott tervezési időszakban alkalmazható üzemmódok az esetek többségében sorbarendezhetők úgy, hogy a megfelelő erőművi egységek sorszámából alkotott $J(i, j)$, hal-

mazokra $i=1, 2, \dots, K, j=1, 2, \dots, M(i)$

$$J(i, j_1) \supset J(i, j_2) \supset \dots \supset J(i, j_{M(i)})$$

teljesüljön.

Az egyszerűsített modellt olyan tervezési időszakra fogalmazzuk meg, amelyben az üzemmódok a fenti tulajdonságúak. Feltételezzük még, hogy az üzemmódok megadásának a sorrendje olyan, hogy

$$J(i, 1) \supset J(i, 2) \supset \dots \supset J(i, M(i))$$

teljesül minden $i=1, 2, \dots, K$ esetén.

Ez a feltételezés úgy is megfogalmazható, hogy üzemmódváltásnál egy nagyobb sorszámú üzemmódra való áttérés berendezéscsoport(ok) kikapcsolásával valósítható meg.

Ezen egyszerűsítő feltevés egyrészt a 4.1. bekezdésben leírtak modellbeli megfogalmazását könnyíti, másrészt az erőművi egységek állásköltségének figyelembevételét egyszerűsíti. Ha ugyanis $J(i, 1) \supset J(i, 2) \supset \dots \supset J(i, M(i))$ teljesül, akkor azok az egységek, amelyekre $k \in J(i, l) \setminus J(i, l+1)$ teljesül, állnak azokban a periódusokban, amelyekben nem az l -edik vagy azt megelőző üzemmódok valamelyike üzemel. Itt és a továbbiakban egy üzemmóddal kapcsolatban használjuk a „ j -edik üzemmódot megelőzi”, ill. „ a j -edik üzemmódot követi” kifejezéseket, ami alatt azt értjük, hogy az üzemmódok felsorolásában a szóbanforgó üzemmód a j -edik előtt, ill. után szerepel, azaz j -nél kisebb, ill. j -nél nagyobb a sorszáma.

Az üzemmódok és erőművi egységek állása közötti fent említett kapcsolat lehetővé teszi, hogy a továbbiakban erőművi egységek állásköltségének figyelembevétele helyett üzemmódok állásköltségét tekintsük.

Legyen $s_{ij}(\tau)$ az i -edik erőmű j -edik üzemmódja állásköltségét megadó függvény. Defináljuk az $s_{ij}(\tau)$ értékét a következő módon:

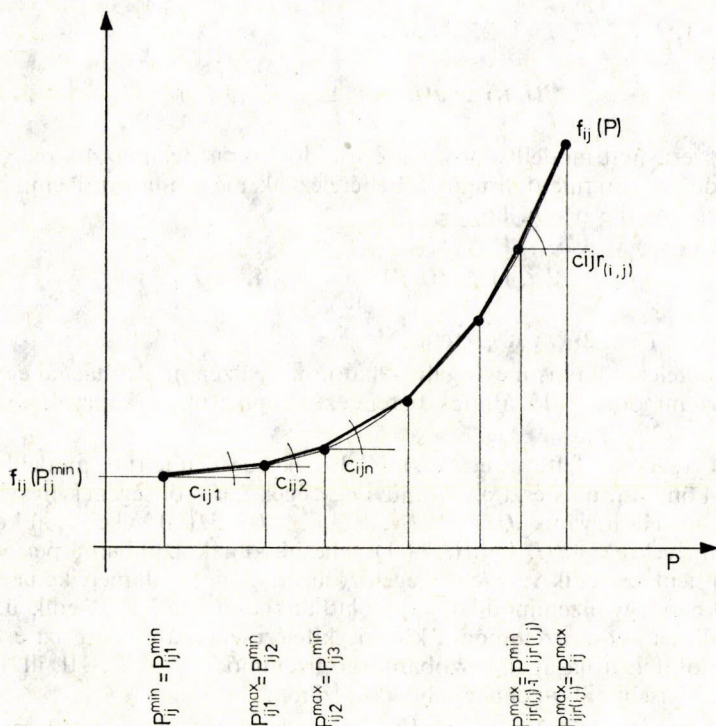
$$s_{ij}(\tau) = \sum_{k \in J(i, j) \setminus J(i, j+1)} g_{ik}(\tau), \quad i = 1, 2, \dots, K, \quad j = 1, 2, \dots, M(i) - 1.$$

A $k \in J(i, j) \setminus J(i, j+1)$ sorszámú erőművi egységek állásköltségét azon periódusokban kell figyelembe vennünk, amelyekben a j -edik üzemmódot követő üzemmódok valamelyike működik, így a j -edik üzemmód $s_{ij}(\tau)$ állásköltségét is azokban a periódusokban kell figyelembe venni, amelyekben a j -edik üzemmódot követő üzemmódok valamelyike működik.

4.3. A termelési költség közelítése

A 3.2.1. bekezdés szerint az i -edik erőmű j -edik üzemmódjának P szintű termelésből származó költségrész az $f_{ij}(P)$ függvényértékkel adható meg, $i=1, 2, \dots, K, j=1, 2, \dots, M(i)$.

Egyszerűsítő feltevésként az $f_{ij}(P)$ nemlineáris költségfüggvények helyett tekintjük ezek lineáris törtfüggvényre való közelítését.



9. ábra

Termelési költség közelítése törtfüggvénnyel

Legyen az i -edik erőmű j -edik üzemmódjához tartozó költségfüggvény közelítő szakaszainak a száma $r(i, j)$. Jelöljük a közelítő függvény töréspontjaihoz tartozó teljesítményértékeket a következő módon:

$$(4.1) \quad P_{ij}^{\min} = P_{ij1}^{\min}, \quad P_{ij1}^{\max} = P_{ij2}^{\min}, \quad P_{ijr(i,j)-1}^{\max} = P_{ijr(i,j)}^{\min}, \quad P_{ijr(i,j)}^{\max} = P_{ij}^{\max},$$

a közelítő szakaszok meredeksége pedig legyen c_{ijl} , $l=1, 2, \dots, r(i, j)$. Az $f_{ij}(P)$ költségfüggvények olyanok, hogy a meredekségekre teljesülnek a következő relációk:

$$(4.2) \quad c_{ij1} < c_{ij2} < \dots < c_{ijr(i,j)}.$$

Az i -edik erőmű j -edik üzemmódja P szintű termelésének költsége e közelítés esetén a következő összeggel adható meg, feltéve, hogy $P_{ijl_0}^{\min} \leq P \leq P_{ijl_0}^{\max}$ (azaz P az l_0 -adik szakaszhoz tartozik):

$$(4.3) \quad f_{ij}(P) \sim K_{ij} + \sum_{l=1}^{l_0-1} (P_{ijl}^{\max} - P_{ijl}^{\min}) c_{ijl} + (P - P_{ijl_0}^{\min}) c_{ijl_0},$$

ahol $K_{ij} = f_{ij}(P_{ij}^{\min})$ a 3. fejezetben bevezetett jelölés szerint.

A termelési költség függvény ilyen szakaszonként lineáris függvénnyel történő helyettesítése esetén a teljesítmény változók számát növelnünk kell. Az általános modell egy p'_{ij} változója helyett az egyszerűsített modellben $r(i, j)$ számú változónk lesz (5.1.2. bekezdés).

4.4. Az üzemmódváltás költségének közelítése

Az üzemmódok átkapcsolásából, állásából származó költségrész megadható a 3.2.2. bekezdésben ismertetett módon. Az ott végeredményként kapott formula azonban bonyolult nemlineáris függvénye az üzemmódokat jellemző 0–1 értékű változóknak.

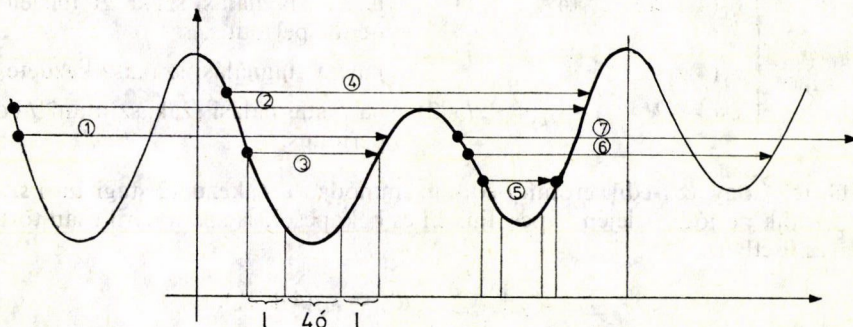
Ebben a bekezdésben megmutatjuk, milyen feltételezéssel lehet ezt a költséget az említett változók lineáris függvényével közelíteni.

A közelítéshez a következő feltevést tesszük: az erőművi egységek kikapcsolás-újraindítása szimmetrikusan történik. Ezen feltevés alatt azt értjük, hogy ha egy erőművi egység a stagnálási szakasz előtt l -periódussal került kikapcsolásra, akkor visszszakapcsolása a stagnálási szakaszt követő l -edik periódus befejeződésekor történik. (Ez a feltételezés szorosan kapcsolódik a 4.1. bekezdésben elmondottakhoz, az igénygörbe jellegzetes tulajdonsága indokolja.) Feltevésünk következményeként nincsenek az egyes tervezési időszakok között ún. „átnyúló” költségek. Ha egy egység a vizsgált időszakot megelőzően kikapcsolásra került, akkor az állásából származó költség közelítő értékét a megelőző időszakban már figyelembe kellett venni. Ehhez hasonlóan, ha egy egység a vizsgált tervezési időszakban kerül kikapcsolásra, akkor ennek költsége a vizsgált időszakban vétetik figyelembe.

A szimmetrikus kikapcsolás-újraindítást csupán az állásidő költség kiszámításának leegyszerűsítésére tételezzük fel, megengedünk azonban olyan menetrendeket is, amelyekben vannak nem szimmetrikus kikapcsolás-újraindítást jelentő üzemmódok is.

Az üzemmódváltás-költség lineáris közelítő függvénye olyan, hogy ha a szimmetrikus kikapcsolás-bekapcsolás érvényesül, akkor a 3.2.2. bekezdésben szereplő költséget szolgáltatja, ettől eltérő esetben ennek közelítő értékét.

A szimmetrikus kikapcsolás-újraindítás fogalmának jobb megértésére tekintünk a 10. ábrát.



10. ábra

Szimmetrikus kikapcsolás-újraindítás esetén az ①, ②, ④, ⑥, ⑦ kikapcsolás-újraindítás nem fordulhat elő

Ezen egymáshoz csatlakozó igénygörbék és kikapcsolás—újraindítás párok találhatók. A nyílak kezdőpontja egy egység kikapcsolását, végpontja az újraindítását jelzi.

	Kikapcsolás időpontja	Bekapcsolás időpontja
1	Előző nap II. kikapcsolási szakasz	I. bekapcsolási szakasz
2	Előző nap II. kikapcsolási szakasz	II. bekapcsolási szakasz
3	I. kikapcsolási szakasz	I. bekapcsolási szakasz
4	I. kikapcsolási szakasz	II. bekapcsolási szakasz
5	II. kikapcsolási szakasz	II. bekapcsolási szakasz
6	II. kikapcsolási szakasz	Következő nap I. bekapcsolási szakasz
7	II. kikapcsolási szakasz	Következő nap II. bekapcsolási szakasz

A szimmetrikus újraindítás azt jelenti, hogy ezek közül csak ③ és ⑤ típusú eset léphet fel, ebből pedig következik, hogy a költségnek nincs a következő napra „át-nyúló” része.

Az ábrán jól látható, hogy a feltételezést az igénygörbe alakjának jellegzetessége indokolja (közelítőleg azonos igényértékek a stagnálási szakaszra „szimmetrikusan” vannak).

Az állásköltség közelítésében egy, a stagnálási szakasz előtt l_0 periódussal kikapcsolt, $4 + 2l_0$ állásidejű egység vagy (az üzemmódváltástól függően) egységek leállítás-újraindítási költségét a 3.2.2. bekezdéshez hasonló módon, összegként fogjuk előállítani (ez az állásidő onnan adódik, hogy a stagnálási szakasz 4 óra, és az azt megelőző és követő l_0 periódus együttes időtartama $2l_0$ óra; itt eltekintünk attól, hogy vannak félórás periódusok is).

Definiáljuk az összegként való előállításához a d_{ij}^t mennyiségeket a következő módon (minden erőmű minden üzemmódjához és minden periódusra, kivéve a stagnálási szakaszok második, harmadik és negyedik periódusát):

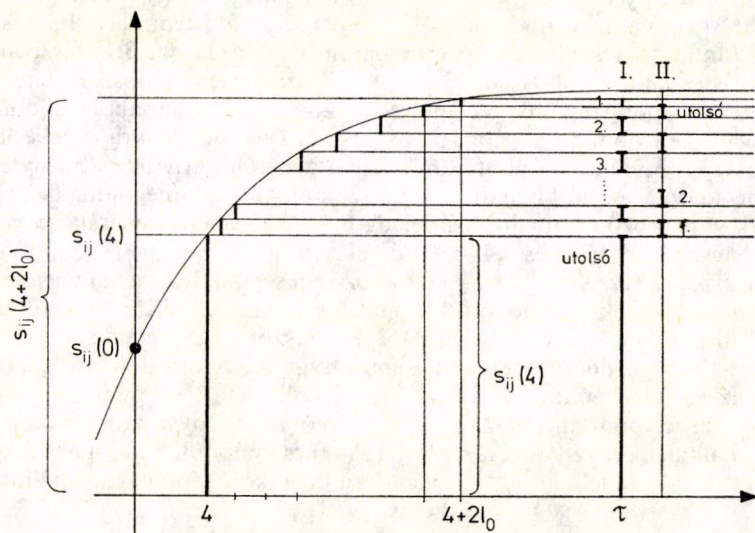
$$(4.4) \quad d_{ij}^t = \begin{cases} s_{ij}(4 + 2l) - s_{ij}(4 + 2l - 1), & \text{ha } t \text{ a stagnálási szakaszt megelőző } l\text{-edik periódus,} \\ s_{ij}(4), & \text{ha } t \text{ a stagnálási szakasz kezdete,} \\ s_{ij}(4 + 2l - 1) - s_{ij}(4 + 2l - 2), & \text{ha } t \text{ stagnálási szakasz utáni } l\text{-edik periódus.} \end{cases}$$

Tegyük fel, hogy az i -edik erőmű j -edik üzemmódját a t_0 kezdetű stagnálási szakasz előtti l_0 -adik periódus elején kapcsoljuk ki és bekapcsolása szimmetrikusan történik. Ebben az esetben:

$$(4.5) \quad \sum_{t=t_0-l_0}^{t_0} d_{ij}^t + \sum_{t=t_0+4}^{3+t_0+l_0} d_{ij}^t = s_{ij}(4 + 2l_0),$$

azaz a d_{ij}^t mennyiségek állásidőre vett összege éppen a megfelelő állásköltség. A d_{ij}^t mennyiségek összegezése ugyanúgy az állásköltség növekményeinek összegezését je-

lenti, mint a 3.2.2. bekezdésben szereplő összegzés. A két összegzés sorrendje azonban eltér egymástól. A 11. ábra ennek megvilágítására szolgál.



11. ábra

Állásköltség előállítás a d_{ij}^* növekmények összegeként.

A jobb oldalon I-vel megjelölt növekmények az összegzés első, a II-vel megjelölt növekmények az összegzés második részét alkotják; az arab számok a tagok sorrendjére utalnak

Amennyiben a bekapcsolás nem szimmetrikusan történik, és az összegzést a tényleges állásidőre végezzük el, az $s_{ij}(4 + 2l_0)$ értéktől az összegzés második tagjához tartozó részletösszeggel térünk el. Feltételezzük, hogy az így kapott állásköltség elég jól közelíti a tényleges $s_{ij}(\tau)$ költséget.

A modellben gondoskodnunk kell arról, hogy a 10. ábra jelöléseiben ①, ill. ② típusú kikapcsolás-újraindítás miatt álló egységek állásköltsége ne terhelje a vizsgált időszakot. A ④ típusú kikapcsolás-újraindítás miatt a második kikapcsolás-stagnálás-bekapcsolás szakaszban is álló berendezésekhez a második részben már ne számoljunk állásköltséget. A ⑥ és ⑦ típusú kikapcsolás-újraindítás költségét a szimmetrikus kikapcsolás-újraindítás költség fentiekben ismertetett közelítésével adjuk meg. A ③ és ⑤ típusú kikapcsolás-újraindítás vagy szimmetrikus, vagy annak közelítése.

4.5. Munkapont bevezetése, néhány jelölés

A 3. fejezetben leírt modellnek az átviteli hálózatra vonatkozó részében a (3.20), (3.22), (3.24) feltételekben szereplő függvények, valamint a (3.10) célfüggvény nemlineáris függvények. Mivel lineáris modell felépítését tűzzük ki célul, ezért ezeket lineáris függvényekkel kívánjuk közelíteni. Eljárásunk az lesz, hogy rögzítünk egy munkapontot, erre vonatkozóan linearizálunk, és a változókra vonatkozó korlátok-

kal biztosítjuk, hogy a munkapont így adódó környezetében a közelítés elegendően pontos legyen.

A munkapont természetesen az (F3. 43) hálózati feltételeket kielégítő pont lesz, így meghatározása a hálózatszámítási feladat megoldását jelenti (lásd F3. rész). Ehhez meg kell adnunk a feszültségtartó csomópontok — ezek a 3.3.1. részben definiált meddőforrás csomópontok (lásd F3. rész) — feszültségeinek abszolút értékét, valamint a referenciapont feszültségét, melyek a modell szempontjából kiindulási, input adatok. Meg kell adnunk továbbá a fogyasztói csomópontokra a teljesítmény-fogyasztást, az erőművi csomópontokon pedig a generált aktív teljesítményt. Az előbbieket input adatok, az utóbbiakat pedig úgy kapjuk, hogy megoldunk egy, az illető periódusra vonatkozó optimalizálási feladatot a teljesítmény-generálásra vonatkozóan. (Részletesen lásd 5.3.1. és 6.4. bekezdéseket.)

A linearizálás részletesebb tárgyalásához szükségünk lesz néhány jelölésre. A továbbiakban a változóknál a periódus azonosítására szolgáló felső t indexet elhagyjuk. Jelöljük a munkapontot $(\mathbf{v}^*, \mathbf{w}^*)$ -gal. A 3.3.1. részben bevezetett jelöléstől eltérően $L(=N_M)$ jelöli a meddőforrás csomópontok számát. A meddőforrás csomópontok halmazát az F3. részben tárgyaltak szerint a feszültségtartó vagy (Q, V) csomópontok és a referenciapont alkotják. A többi csomópontot fogyasztói csomópontoknak nevezzük, halmazukat az F3. rész szerinti (P, Q) csomópontok alkotják. A fogyasztói csomópontok számát jelölje N_F . A \mathbf{v} vektort a következőképp particionáljuk:

$$(4.6) \quad \mathbf{v} = \begin{pmatrix} \mathbf{v}^M \\ \mathbf{v}^F \end{pmatrix},$$

ahol $\mathbf{v}^M \in R^L$ a meddőforrás csomópontokhoz, $\mathbf{v}^F \in R^{N_F}$ pedig a fogyasztói csomópontokhoz tartozik, és feltételeztük, hogy a csomópontok sorszámozása a meddőforrásokkal kezdődik. Felhívjuk az Olvasó figyelmét, hogy az F felső indexet kétféle értelemben is használjuk: feszültséggel kapcsolatban a fentiek szerint fogyasztói csomópontokat, míg a hatásos teljesítmény vonatkozásában teljesítmény-fogyasztást (igényt) azonosít.

Legyen $\mathbf{g}^T(\mathbf{v}, \mathbf{w}) = (g_1(\mathbf{v}, \mathbf{w}), \dots, g_N(\mathbf{v}, \mathbf{w}))$, ezt is particionáljuk a meddőforrás és a fogyasztói komponensek szerint:

$$(4.7) \quad \begin{aligned} (\mathbf{g}^M(\mathbf{v}, \mathbf{w}))^T &= (g_1(\mathbf{v}, \mathbf{w}), \dots, g_L(\mathbf{v}, \mathbf{w})), \\ (\mathbf{g}^F(\mathbf{v}, \mathbf{w}))^T &= (g_{L+1}(\mathbf{v}, \mathbf{w}), \dots, g_N(\mathbf{v}, \mathbf{w})). \end{aligned}$$

Hasonlóképpen particionálva a (3.15) összefüggésben a jobb oldalon $\mathbf{Q} \in R^N$ -et:

$$(4.8) \quad (\mathbf{Q}^M)^T = (Q_1, \dots, Q_L); (\mathbf{Q}^F)^T = (Q_{L+1}, \dots, Q_N).$$

Jelöléseinkkel a (3.15) összefüggéseknek a reaktív teljesítményre vonatkozó része a következő alakot ölti:

$$(4.9) \quad \mathbf{g}(\mathbf{v}, \mathbf{w}) = \mathbf{Q}, \quad \text{ill.} \quad \begin{aligned} \mathbf{g}^M(\mathbf{v}, \mathbf{w}) &= \mathbf{Q}^M, \\ \mathbf{g}^F(\mathbf{v}, \mathbf{w}) &= \mathbf{Q}^F. \end{aligned}$$

Szükségünk lesz a $\mathbf{g}(\mathbf{v}, \mathbf{w}^*): R^N \rightarrow R^N$ leképezés *Jacobi-mátrixára* a $\mathbf{v} = \mathbf{v}^*$ pontban, jelöljük azt \mathbf{Y} -nal és particionáljuk a meddő és fogyasztói csomópontoknak meg-

felelően:

$$(4.10) \quad \mathbf{Y} = \underbrace{\begin{pmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 \\ \mathbf{Y}_3 & \mathbf{Y}_4 \end{pmatrix}}_L \underbrace{\}_{N_F}}_{N_F}^L$$

(F3. 35), (F3. 39) alapján az \mathbf{Y} mátrix elemei a következők lesznek:

$$(4.11) \quad Y_{ii} = \sum_{k \in J(i)} w_k^* G_{ik} + v_i^* \sum_{k \in J(i)} (2B_{ik} - \omega C_{ik}) - \sum_{k \in J(i)} v_k^* B_{ik}, \quad i = 1, \dots, N.$$

$$(4.12) \quad Y_{ik} = \begin{cases} -w_i^* G_{ik} - v_i^* B_{ik}; & \text{ha } (i, k) \text{ ága a hálózathoz,} \\ 0, & \text{különben.} \end{cases}$$

4.6. A feszültségi változók számának redukciója

A 3. fejezet modelljében valamennyi csomópont komplex feszültségének valós és képzetes része változóként szerepel. Ebben a részben közelítő lineáris összefüggést konstruálunk \mathbf{v}^M és \mathbf{v}^F között, melynek segítségével \mathbf{v}^F kifejezhető lesz \mathbf{v}^M függvényeként. Ily módon az egyszerűsített modellben elegendő lesz \mathbf{v}^M -et változónak tekinteni.

Az F4. részben megmutatjuk, hogy a reaktívteljesítmény-áramlás elsősorban a csomóponti feszültségek abszolút értékétől függ. A feszültségnek valós és képzetes részre történő felbontását tekintve tapasztalati tény, hogy a reaktívteljesítmény-áramlás elsősorban a valós résztől függ és csak kisebb mértékben a képzetes résztől. Az F4. rész gondolatmeneteihez hasonlóan ez a tény heurisztikus érvekkel is alátámasztható. (4.9)-et tekintve ezért lerögzítjük a képzetes részt a munkaponti w^* értékre. Ily módon a következő összefüggés adódik:

$$(4.13) \quad \mathbf{g}^F(\mathbf{v}^M, \mathbf{v}^F, \mathbf{w}^*) = \mathbf{Q}^F.$$

Az egyenletrendszer *Jacobi-mátrixa* (4.10) alapján $(\mathbf{Y}_3, \mathbf{Y}_4)$. A villamos hálózat (mint gráf) összefüggőségi viszonyaiból következően \mathbf{Y}_4 nem szinguláris. Ezért $(\mathbf{v}^{*M}, \mathbf{v}^{*F})$ egy környezetében a (4.13) egyenletrendszer implicit módon definiálja \mathbf{v}^F -et, mint \mathbf{v}^M függvényét, azaz van olyan $\mathbf{h}: R^M \rightarrow R^{N_F}$ leképezés, hogy $\mathbf{v}^F = \mathbf{h}(\mathbf{v}^M)$. Az implicit függvények deriváltjára vonatkozó tétel szerint ekkor:

$$(4.14) \quad \frac{d\mathbf{h}(\mathbf{v}^{*M})}{d\mathbf{v}^M} = -\mathbf{Y}_4^{-1} \mathbf{Y}_3.$$

A $\mathbf{h}(\mathbf{v}^M)$ függvény \mathbf{v}^{*M} körüli *Taylor-sorában* elhanyagolva a magasabbrendű tagokat, a következő lineáris összefüggés érvényes közelítőleg, a \mathbf{v}^{*M} elegendően kicsi környezetében:

$$(4.15) \quad \mathbf{v}^F - \mathbf{v}^{*F} \sim -\mathbf{Y}_4^{-1} \mathbf{Y}_3 (\mathbf{v}^M - \mathbf{v}^{*M}).$$

A közelítőleg érvényes egyenlőséget itt és a továbbiakban is „ \sim ” jelöli.

4.7. *A feszültségek képzetes részének kifejezése a hatásosteljesítmény-betáplálással
A feszültségi változók számának további redukciója*

Ebben a részben megmutatjuk, hogy linearizálás révén a csomóponti feszültségek képzetes részei kifejezhetők a csomóponti generált \mathbf{P}^G teljesítménnyel. Mivel ez utóbbi a felépítendő modell változóival lineárisan kifejezhető lesz, ezért lehetőség nyílik arra, hogy az egyszerűsített modellben w ne szerepeljen változóként.

Itt is kínálkozik a 4.6. bekezdésben alkalmazott, az implicit függvényekre épített linearizálás. Most mégis másképp fogunk eljárni, olyan eljárást alkalmazunk, amely az ágterhelési feltételekhez egyébként is szükséges teljesítményáramlási összefüggések linearizálására épül.

Ismeretes, hogy a hatásosteljesítmény-áramlás elsősorban a fázisszögekre érzékeny. A feszültségek valós és képzetes részre bontása esetén hasonlóan megmutatható, hogy a hatásosteljesítmény-áramlás a feszültségek képzetes részére érzékeny elsősorban. Ezért lerögzítjük a feszültség valós részét a munkaponti v^* értékre, és a teljesítményáramlást w^* kicsi környezetében mint w függvényét vizsgáljuk. A linearizáláshoz reprodukáljuk az i -edik csomópontból az (i, k) ágra kiáramló hatásos teljesítményre vonatkozó (F4. 32) összefüggést:

$$(4.16) \quad T_{ik}(v_i^*, v_k^*, w_i, w_k) = G_{ik}[v_i^*(v_i^* - v_k^*) + w_i(w_i - w_k)] + B_{ik}[w_i v_k^* - w_k v_i^*].$$

Az F4. részben foglaltak szerint $G_{ik} \ll B_{ik}$, továbbá viszonylagos egységekben (lásd F4. rész) v_i^* és v_k^* , valamint w_i és w_k eltérése kicsi, ezért a gyakorlatban jól bevált alábbi közelítéssel élünk:

$$(4.17) \quad T_{ik}(v_i^*, v_k^*, w_i, w_k) \sim B_{ik}(w_i v_k^* - w_k v_i^*).$$

A kapott összefüggést valamennyi ágra felírjuk, ami kompakt módon tehető meg a következő jelölések felhasználásával.

Legyen V egy $M \times N$ méretű mátrix, ami az A csúcs, él incidencia-mátrix A^T transzformáltjából úgy adódik, hogy abban soronként haladva az ág kezdőpontja pozíciójába a munkapontból adódó végponti feszültség valós részét, míg a végpont helyére a kezdőponti feszültség valós részének (-1) -szeresét írjuk. Jelöljük az ágak kezdőpontjából az ágra áramló hatásos teljesítmények vektorát T -vel, $T \in R^M$. B legyen $M \times M$ méretű diagonális mátrix; ha az l -edik ág (i, k) , akkor az átló l -edik eleme legyen B_{ik} .

Jelöléseinkkel a teljesítményáramlásra vonatkozó közelítés:

$$(4.18) \quad T \sim BVw.$$

Megjegyezzük, hogy a (4.18) összefüggések a gráf irányítása által definiált referencia-irányra vonatkoznak áganként, azaz az irányított gráfban az ágak kezdőpontjából az ágra kimenő teljesítményt írják le. A kimenő teljesítményt természetesen úgy értjük, hogy ha a tényleges teljesítményáramlás az (i, k) ág esetén $k \rightarrow i$ irányú, akkor a kimenő teljesítmény i -ből az ágra negatív értékű.

A 3.3.1. bekezdésnek megfelelően itt is jelöljük \mathbf{P}^G -vel ($\mathbf{P}^G \in R^N$) a csomópontként generált teljesítmények vektorát, ekkor persze $P_i^G = 0$, $i > L$ -re; \mathbf{P}^F a csomópontkénti fogyasztások vektora. Ha valamely mennyiséget $*$ -gal is megjelölünk, az a munkapontra fog vonatkozni.

A T teljesítményáramlást a \mathbf{P}^G vektorral összekapcsoló összefüggésként természetes módon kínálkozik Kirchhoff csomóponti törvényének megfelelője (lásd F4. rész)

a teljesítményáramlásra vonatkozóan. Mivel a T teljesítményáramlás a referencia-irányokra vonatkozik, és az ágakon teljesítményveszteség lép fel, ezért a *Kirchhoff törvényben* korrekciós tagot is be kell vezetni. Az eljárás egyszerűen az, hogy ha a tényleges áramlás az irányításnak megfelelő, akkor a végponti fogyasztást csökkentjük a veszteséggel, ellenkező esetben ugyanazzal növeljük.

Legyen $P^K \in R^N$ a következő vektor (a jelölésre lásd F4.):

$$(4.19) \quad P_l^K = - \sum_{\substack{(i,l) \text{ ág} \\ \text{sgn } T_{il} \geq 0}} P_{il}^v + \sum_{\substack{(i,l) \text{ ág} \\ \text{sgn } T_{il} < 0}} P_{il}^v, \quad l = 1, \dots, N.$$

Ezek után *Kirchhoff csomóponti törvénye* a munkapontban:

$$(4.20) \quad P^{*G} = P^F + P^{*K} + A^T T^*.$$

Itt P^{*K} a munkaponti feszültségeloszlás ismeretében (F4. 22), ill. (F4. 30) felhasználásával, (4.19) alapján meghatározható. Rögzítve a munkaponthoz tartozó veszteségeket, a munkapont kis környezetére a következő közelítés adódik:

$$(4.21) \quad P^G \sim P^F + P^{*K} + A^T T.$$

Behelyettesítjük ide a (4.18) összefüggésből a teljesítményáramlásra vonatkozó közelítést, így az alábbi összefüggés adódik:

$$(4.22) \quad P^G - P^F - P^{*K} \sim A^T B V w.$$

Itt az $A^T B V$ mátrix $N \times N$ méretű, a rangja $N-1$ (lásd F1. 8 tétel).

Töröljük $A^T B V$ mátrix referenciapontnak megfelelő sorát és oszlopát, a maradt mátrix már nonsinguláris lesz (lásd F1. 8 tétel). Invertáljuk, majd az inverzet kiegészítjük egy csupa 0 első sorral és oszloppal, a kapott mátrixot jelölje Z^N . A referenciapontban a potenciál fázisszöge 0, ezért $w_1 = 0$. Így a (4.22) összefüggésből adódik a címben ígért közelítő formula:

$$(4.23) \quad w \sim Z^N (P^G - P^F - P^{*K}).$$

4.8. A hálózati feltételek linearizálása

A 4.6. és 4.7. bekezdésekben tárgyaltak alapján adott olyan egyszerűsített lineáris modell felépítésének a lehetősége, melyben a hálózathoz kapcsolódó feltételrendszerben teljesítmény és üzemállapot típusú változók mellett csak a meddőforrás csomópontok komplex feszültségének valós részei szerepelnek változóként. Ehhez azonban egyrészt meg kell adnunk az általános modell (3.22), (3.24) feltételrendszereiben szereplő függvények megfelelő lineáris közelítéseit, másrészt a (3.20) hálózati feltételrendszer kezelésmódját.

Először a (3.22) ácterhelési feltételrendszert vizsgáljuk. A linearizálás (4.18) és (4.23) felhasználásával könnyen adódik:

$$(4.24) \quad T \sim B V Z^N (P^G - P^F - P^{*K}).$$

Itt a jobb oldalon csupán P^G változó, a későbbiek során megmutatjuk, hogy az egyszerűsített modell teljesítmény és üzemállapot változóival lineárisan kifejezhető (lásd 5.1.3. bekezdés).

Az ágterhelési feltételekkel kapcsolatban egy további elhanyagolással is fogunk élni, amennyiben a (3.22) feltételekben szereplő $|S_{ik}|$ látszólagos teljesítményt a veszteséggel korrigált hatásos teljesítmény abszolút értékével helyettesítjük, megfelelően korrigált T_{ik}^{\max} korlátokat használva.

Ezután a (3.24) meddőforrás feltételrendszert tekintjük. A 4.6. bekezdés jelölésrendszerét használva a $\mathbf{g}^M(\mathbf{v}^M, \mathbf{v}^F, \mathbf{w}^*)$ -nak a $(\mathbf{v}^{*M}, \mathbf{v}^{*F})$ körüli *Taylor sorfejtésének* kvadratikus tagjait elhanyagolva adódik a következő:

$$(4.25) \quad \mathbf{g}^M(\mathbf{v}^M, \mathbf{v}^F, \mathbf{w}^*) \sim \mathbf{g}^M(\mathbf{v}^{*M}, \mathbf{v}^{*F}, \mathbf{w}^*) + \mathbf{Y}_1(\mathbf{v}^M - \mathbf{v}^{*M}) + \mathbf{Y}_2(\mathbf{v}^F - \mathbf{v}^{*F}).$$

Behelyettesítünk ide (4.15) alapján:

$$(4.26) \quad \mathbf{g}^M(\mathbf{v}^M, \mathbf{v}^F, \mathbf{w}^*) \sim \mathbf{g}^M(\mathbf{v}^{*M}, \mathbf{v}^{*F}, \mathbf{w}^*) + (\mathbf{Y}_1 - \mathbf{Y}_2 \mathbf{Y}_4^{-1} \mathbf{Y}_3)(\mathbf{v}^M - \mathbf{v}^{*M}).$$

Végül a (3.20) hálózati feltételrendszerrel kapcsolatban teszünk néhány megjegyzést. Ez kvadratikus, egyenlőség alakú feltételekből áll, melyeket a hálózati feszültség-eloszlásnak ki kell elégíteni. A modellbe implicit módon épülnek be, a fentiekben tárgyalt linearizálásokból adódó feltételeken keresztül, a linearizálási eljárás hibájának még megengedett értéke által meghatározott tartományban. A napi menetrend optimalizálására vonatkozó modellt megoldjuk az egyenlőtlenséges hálózati feltételekkel, a meddőforrás csomópontok feszültsége valós részét tekintve változóknak. A megoldáshoz hozzátéve a meddőforrás csomóponti feszültségek munkaponthoz tartozó képzetes részét, az optimális generált teljesítményeket alapul véve, megoldjuk a hálózatszámítási feladatot. (Lásd F3., F4. részek). Amennyiben a csomóponti feszültségekre vonatkozó alsó-felső korlátok a munkaponti értékekhez elég közeliek, azaz a tartomány elég szűk, akkor az ilymódon kapott feszültség és teljesítmény-eloszlás fizikailag továbbra is kielégíti a feszültségtartási, ágterhelési és meddőforrás feltételeket. A stratégia a gyakorlatban jól bevált.

4.9. A hálózati veszteség függvény linearizálása

A $C^v(\mathbf{v}, \mathbf{w})$ hálózati veszteség függvény, (3.9), a (\mathbf{v}, \mathbf{w}) változók kvadratikus, konvex függvénye. A munkapont körüli linearizálás egyszerűen azt jelenti, hogy a $(\mathbf{v}^*, \mathbf{w}^*)$ pont körüli *Taylor-sorfejtésből* elhanyagoljuk a kvadratikus tagokat.

A 4.6. bekezdés jelölésrendszerét használva, a munkaponti értékeket $*$ -gal jelölve, a következő adódik:

$$(4.27) \quad P^v(\mathbf{v}, \mathbf{w}) \sim P^v(\mathbf{v}^*, \mathbf{w}^*) + \nabla_{\mathbf{v}}^T P^v(\mathbf{v}^*, \mathbf{w}^*)(\mathbf{v}^M - \mathbf{v}^{*M}) + \\ - \nabla_{\mathbf{v}^F}^T P^v(\mathbf{v}^*, \mathbf{w}^*)(\mathbf{v}^F - \mathbf{v}^{*F}) + \nabla_{\mathbf{w}}^T P^v(\mathbf{v}^*, \mathbf{w}^*)(\mathbf{w} - \mathbf{w}^*).$$

A tapasztalatok szerint itt a jobb oldalon a harmadik tag elhanyagolható. Bevezetjük a következő jelöléseket: $\mathbf{d}^v = \nabla_{\mathbf{v}}^T P^v(\mathbf{v}^*, \mathbf{w}^*)$, $\mathbf{b}^v = \nabla_{\mathbf{w}}^T P^v(\mathbf{w}^*, \mathbf{w}^*)$. Az (F4. 30) egyenlősből közvetlen differenciálással adódik a következő:

$$(4.28) \quad d_i^v = \sum_{\substack{k \\ (i, k) \text{ ág}}} R_{ik} \left[2G_{ik} I_{ik}^{*P} - \left(B_{ik} - \frac{1}{2} \omega C_{ik} \right) I_{ik}^{*Q} + B_{ik} I_{ki}^{*Q} \right], \quad i = 1, \dots, L.$$

Hasonlóképpen kaphatók \mathbf{b}^v komponensei:

$$b_i^v = \sum_{\substack{k \\ (i,k) \text{ ág}}} R_{ik} [2B_{ik} I_{ki}^{*P} + G_{ik} I_{ik}^{*Q} - G_{ik} I_{ik}^{*Q}], \quad i = 1, \dots, N.$$

Felhasználva még a (4.23) összefüggést, a (4.27) a következőképp írható:

$$(4.29) \quad P^v(\mathbf{v}, \mathbf{w}) \sim P^v(\mathbf{v}^*, \mathbf{w}^*) + \sum_{i=1}^L d_i^v v_i - \sum_{i=1}^L d_i^v v_i^* + \mathbf{b}^{vT} \mathbf{Z}^N (\mathbf{P}^G - \mathbf{P}^F - \mathbf{P}^{*K}) - \mathbf{b}^{vT} \mathbf{w}^*.$$

A későbbiek során megkonstruált modellben \mathbf{P}^G kifejezhető lesz lineárisan a teljesítmény és üzemállapot típusú változókkal, amint azt már korábban is említettük.

4.10. Feszültségellenőrző periódusok

Feszültségellenőrző periódusoknak azokat a periódusokat nevezzük, amelyeknél a modellben a kalorikus feltételek mellett hálózati feltételeket is előírunk. A hálózati feltételrendszer generálása és az optimalizálásban való figyelembevétele viszonylag nagyobb számítási-idő igényű, ezért az egyszerűsített modellben csupán három periódusra írunk elő hálózati feltételeket. Ezek a szélsőséges terhelésű periódusok: az esti csúcsterhelési periódus, az éjszakai minimum és a délelőtti csúcs (lásd 4.1.). Csúcsterhelésnél nagy a meddőteljesítmény-fogyasztás és alacsonyabbak a feszültségek, az éjszakai minimumban pedig nagy a meddőteljesítmény-termelés, magasak a feszültségek. Gyakorlati tapasztalat, és a 4.1.-ben megfogalmazott erőművi blokkokra vonatkozó be- és kikapcsolási stratégiából is következtethető, hogy amennyiben a szélsőséges terhelésű periódusokban megköveteljük a hálózati feltételek teljesülését, azok a közbülső, ún. normál periódusokban is teljesülni fognak.

4.11. A hálózati feltételek kiválasztása

A (3.21) feszültségtartási feltételek valamennyi csomópontra vonatkoznak. Ezek közül a képzetes részt korlátozó feltételeket elhagyjuk az egyszerűsített modellben, ugyanis (4.23)-ból láthatóan a hatásos teljesítményt korlátozva a képzetes részt is korlátozzuk, másrészt tapasztalati tény, hogy a képzetes rész változása a munkapont környezetében kicsi a többi villamos mennyiségekre előírt szigorú korlátok következtében.

A csomóponti feszültségek valós részeire előírt korlátok esetében a fogyasztói csomópontokat tekintve ezek közül csak néhány olyan van, amelyik a feszültségtartás szempontjából kritikus.

A munkapontot tekintve a fogyasztói csomópontokra vonatkozó feltételek közül kiválasztjuk a K_v számú legjobban megsértett feltételt, ill. kielégítettség esetén azokat, ahol a munkaponti értékek a legközelebb vannak a határhoz. A gyakorlatban $K_v=30$ elegendőnek bizonyult.

A (3.22) ágrterhelési feltételekben az összes vezetéknek csupán kis hányada az, amely túlterhelés szempontjából szóba kerül. Így a munkaponti teljesítményáramlást alapul véve a feszültségtartási feltételekhez analóg módon kiválasztunk K_A darab feltételt, és csupán ezeket írjuk elő az egyszerűsített modellben. A gyakorlatban itt is elegendő a $K_A=30$ érték.

5. Az egyszerűsítő feltételek figyelembevételével kapott modell

A 4. fejezetben leírtak alapján az ütemezési probléma általános modellje egyszerűsíthető.

Az egyszerűsített modellnek nagyméretű, vegyesváltozós, matematikai programozási feladat felel meg, melynek célfüggvénye és feltételei is lineárisak, együttható mátrixa strukturált. Megoldására van lehetőség.

Külön megemlítjük, hogy a modell felépítésénél az említett egyszerűsítéseken túl egy további jelölésbeli változtatást is teszünk: az egyszerűsített modellben az üzemmód változókat az általános modelltől eltérően definiáljuk, hogy a modell 0–1 értékű változóinak a számát csökkentsük.

E modell leírásakor is alkalmazkodunk a 3.1. bekezdésben ismertetett jelölés-technikai konvenciókhoz.

5.1. Az egyszerűsített modell

5.1.1. Üzem mód változók

Definiáljuk az x_{ij}^t , $i=1, 2, \dots, K$; $j=1, 2, \dots, M(i)-1$; $t=1, 2, \dots, 27$ változókat a következő módon (K az erőművek számát, $M(i)$ az i -edik erőműben alkalmazható üzemmódok számát jelöli, ugyanúgy, mint az általános modell leírásánál):
Legyen

$$(5.1) \quad x_{ij}^t = \begin{cases} 0, & \text{ha az } i\text{-edik erőműben a } t\text{-edik periódusban a } j\text{-edik vagy} \\ & \text{a } j\text{-ediket megelőző üzemmódok valamelyike működik;} \\ 1, & \text{ha az } i\text{-edik erőműben a } t\text{-edik periódusban valamely,} \\ & \text{a } j\text{-ediket követő üzemmód üzemel.} \end{cases}$$

Világos, hogy ha az $x_{i,M(i)}^t$, $i=1, 2, \dots, K$; $t=1, 2, \dots, 27$ szimbólumok jelentését is a fentiek szerint definiáljuk, ennek értéke csak 0 lehet. Így $x_{i,M(i)}^t=0$, $i=1, 2, \dots, K$; $t=1, 2, \dots, 27$, és ezek nem változói a modellnek, az egy-egy erőműhöz rendeltendő üzemmód változók száma csak $M(i)-1$.

A modell leírásában ennek ellenére előfordulnak az $x_{i,M(i)}^t$ szimbólumok, az említett értelemben. Ehhez hasonlóan használjuk az $x_{i,0}^t$, $i=1, 2, \dots, K$; $t=1, 2, \dots, 27$ jeleket. Ezek értéke definíció szerint 1.

Ilyen definíció esetén az $x_{i,j-1}^t - x_{ij}^t$ különbség értéke akkor és csak akkor 1, ha a t -edik periódusban, az i -edik erőműben a j -edik üzemmód üzemel ($i=1, 2, \dots, K$, $j=1, 2, \dots, M(i)$, $t=1, 2, \dots, 27$). x_{ij}^t definíciójából az is következik, hogy az adott erőműhöz tartozó komponensek — azaz az $x_{i1}^t, x_{i2}^t, \dots, x_{i,M(i)-1}^t$ értékek — egymást követően 1 értékeket, és ezekhez csatlakozva egymást követően csak 0 értékeket vehetnek fel. Az egymás mellett álló, 1, 0 párban a 0 érték helye jelzi, hogy hányadik üzemmód üzemel az adott erőműben, az adott periódusban.

Különböző periódusokhoz tartozó fenti változócsoporthoz az 1, 0 pár helye más és más lehet. Jobbrtolódása azt jelenti, hogy az erőműben a periódusok közötti időszakban kikapcsolás történt. Balrattolódása bekapcsolásra utal. Ha az 1, 0 pár helye változatlan, akkor a periódusok közötti időszakban az üzemmód nem változott (lásd 4.2. bekezdés).

Az x'_{ij} , $i=1, 2, \dots, K$, $j=1, 2, \dots, M(i)-1$ komponenseket az erőművek sorrendjében, azon belül pedig az üzemmódok sorrendjében egyetlen vektorra egybekapcsolva kapjuk az x' , t -edik periódushoz tartozó üzemmódváltozó vektort. Ennek dimenziója $\sum_{i=1}^K (M(i)-1)$.

A 4.1. bekezdésben szereplő feltételezés miatt stagnálási szakaszban nem történhet üzemmódváltogatás, elegendő azért a stagnálási szakasz egyetlen, első periódusához tartozó üzemmód változókat tekinteni a modell változóinak. Ezért x' csak $t=1, 2, \dots, t_0, t_0+4, \dots, t_1, t_1+4, \dots, 27$ esetén jelöl változókat, ahol t_0 és t_1 a stagnálási szakaszok első periódusainak sorszáma. Az x felső index nélküli változó ezen x' változók együttesét jelöli. x dimenziója

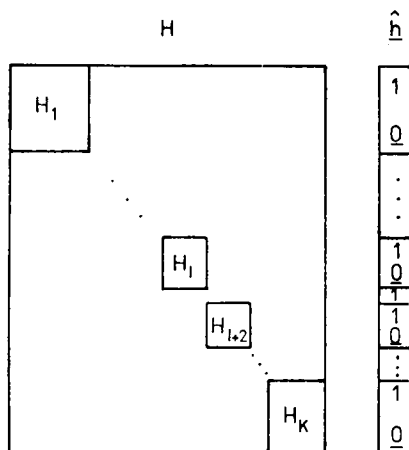
$$21 \sum_{i=1}^K (M(i)-1).$$

A modell leírásában előfordul az x'_{ij} szimbólum $t=t_0+1, t_0+2, t_0+3$ és t_1+1, t_1+2, t_1+3 esetén is, helyettük x'_{ij} , ill. x''_{ij} értendő. A modellben szükséges a tervezési időszakot megelőző időszak legutolsó periódusában alkalmazott üzemmódok ismerete. Az erre vonatkozó információ megadható az ezen legutolsó periódushoz tartozó üzemmód változó értékének megadásával. Jelöljük x^0 -val a megfelelő üzemmód vektort. Ez a modellben $\sum_{i=1}^K (M(i)-1)$ dimenziós, 0 és 1 értékekből álló konstans vektor.

Megjegyezzük, hogy abban az esetben, ha az általános modellben az üzemmódok és a sorrendjük megegyeznek az e modellbeli üzemmódokkal és sorrendezéssel, akkor a két modell üzemmód változói a következő kapcsolatban vannak egymással:

$$(5.2) \quad y' = Hx' + \hat{h}.$$

A túlságosan komplikált jelölésmód elkerülése céljából H , ill. \hat{h} explicit megadása helyett az alábbi, 12. ábrán szemléltettjük ezeket.



12. ábra

Az ábrán a H_l blokk $M(l) \times (M(l)-1)$ méretű, és $(H_l)_{kk} = -1$, $k=1, 2, \dots, M(l)-1$; $(H_l)_{k+1,k} = 1$, $k=1, 2, \dots, M(l)-1$; a mátrix többi eleme 0. Azon erőműveknek megfelelő sorok, ahol egyetlen üzemmód van, 0-val vannak kitöltve, az ábrán ilyen sor található a H_l és H_{l+2} blokkok között, a \hat{h} vektorban a megfelelő komponens 1.

5.1.2. Teljesítmény változók

Az egyes periódusokban az erőművekben alkalmazandó üzemmódok teljesítményszintjének megadására minden üzemmódhoz minden periódusban $r(i, j)$ változót rendelünk, ahol az $r(i, j)$ az üzemmódhoz tartozó költségfüggvény közelítő szakaszainak a száma.

Jelölje az i -edik erőmű j -edik üzemmódjához rendelt változókat p_{ijl}^t , $l=1, 2, \dots, r(i, j)$, $t=1, 2, \dots, 27$.

Legyen $p_{ijl}^t=0$, $l=1, 2, \dots, r(i, j)$, ha a t -edik periódusban az i -edik erőműnek nem a j -edik üzemmódja üzemel. Egyébként az i -edik erőmű j -edik üzemmódja t -edik periódusbeli P szintű üzemelése esetén $p_{ijl}^t \geq 0$, $l=1, 2, \dots, r(i, j)$, és

$$(5.3) \quad P_{ijk}^{\min} \leq P \leq P_{ijk}^{\max}$$

teljesülésekor legyen

$$(5.4) \quad p_{ijl}^t = P_{ijl}^{\max} - P_{ijl}^{\min}, \quad l = 1, 2, \dots, k-1,$$

$$(5.5) \quad p_{ijk}^t = P - P_{ijk}^{\min},$$

$$(5.6) \quad p_{ijl}^t = 0, \quad l = k+1, \dots, r(i, j).$$

A definícióból következik, hogy csak akkor lehet $p_{ijk}^t > 0$ valamely k index esetén, ha (5.4) teljesül. A definícióból az is következik, hogy

$$(5.7) \quad 0 \leq p_{ijl}^t \leq (P_{ijl}^{\max} - P_{ijl}^{\min})(x_{ij-1}^t - x_{ij}^t),$$

$$i = 1, 2, \dots, K; \quad j = 1, 2, \dots, M(i); \quad l = 1, 2, \dots, r(i, j)$$

$$t = 1, 2, \dots, 27$$

feltételeknek is teljesülniük kell. Ezekben a jobb oldalon szereplő $(x_{ij-1}^t - x_{ij}^t)$ tényező biztosítja azt, hogy $p_{ijl}^t > 0$ csak abban az esetben lehetséges, ha az i -edik erőműben a j -edik üzemmód üzemel.

Az így definiált teljesítmény változókkal az i -edik erőmű j -edik üzemmódjának t -edik periódusbeli teljesítményszintje a

$$(5.8) \quad P = P_{ij}^{\min}(x_{ij-1}^t - x_{ij}^t) + \sum_{l=1}^{r(i,j)} p_{ijl}^t$$

összeggel adható meg. Az i -edik erőmű összes üzemmódjaira összegezve ezeket, az erőmű t -edik periódusbeli teljesítményszintjét kapjuk:

$$(5.9) \quad \sum_{j=1}^{M(i)} \left[P_{ij}^{\min}(x_{i,j-1}^t - x_{ij}^t) + \sum_{l=1}^{r(i,j)} p_{ijl}^t \right].$$

Rögzített t esetén jelölje \mathbf{p}^t azt a vektort, amelyet a p_{ijl}^t komponensekből alkotunk úgy, hogy erőművenként és üzemmódonként felsoroljuk a hozzájuk rendelt teljesítmény változókat.

Ez a \mathbf{p}^t a modell t -edik periódusához tartozó teljesítményváltozó vektora. Dimenziója

$$\sum_{i=1}^K \sum_{j=1}^{M(i)} r(i, j).$$

Felső index nélküli \mathbf{p} vektorral a modell teljesítményváltozó vektorát jelöljük. Ez a \mathbf{p}^t vektorok egymás után fűzésével jön létre, $t=1, 2, \dots, 27$.

5.1.3. Feszültségi változók

A feszültségellenőrző periódusokban a teljesítmény és üzemállapot mellett a feszültségtartó csomópontok feszültségének valós részei is változóként szerepelnek. Az ezekből alkotott L -dimenziós vektort a továbbiakban \mathbf{u} -val jelöljük.

A 4. fejezetben feltételeztük, hogy a csomópontonkénti generált hatásos teljesítmények \mathbf{P}^G vektora kifejezhető lineárisan a modell \mathbf{p} , \mathbf{x} változóival. Valóban, legyen \mathbf{D}^* a következő, $N \times N_v$ méretű mátrix (ahol N_v -vel jelöltük a $(\mathbf{p}^T, \mathbf{x}^T)$ dimenzióját): 0-tól különböző elemei csak azon csomópontoknak megfelelő sorokban vannak, amelyekhez erőmű csatlakozik. E sorokban a csatlakozó erőművekhez tartozó teljesítmény-típusú változók oszlopában 1 van, az üzemállapot-típusú változók oszlopai-ban pedig a megfelelő $P_{i,j+1}^{\min} - P_{ij}^{\min}$ mennyiségek. Könnyen látható, hogy

$$(5.10) \quad \mathbf{P}^G = \mathbf{D}^* \begin{pmatrix} \mathbf{p} \\ \mathbf{x} \end{pmatrix} + \mathbf{P}^{\min},$$

ahol \mathbf{P}_k^{\min} a k -adik csomópont-hoz tartozó erőművek üzemmódjainak minimális \mathbf{P}_k^{\min} teljesítményszintjeinek összege, $k=1, 2, \dots, N$ és $\mathbf{P}_k^{\min}=0$, ha a k -adik csomópont-hoz nem csatlakozik erőmű.

5.2. A célfüggvény

A 4. fejezetben ismertetett egyszerűsítő feltevések figyelembevételével és az egyszerűsített modell változóinak felhasználásával a 3.2. bekezdésben leírt nemlineáris célfüggvény összetevői a következőképpen módosulnak:

5.2.1. Erőművi blokk termelési költsége

Az erőművek üzemmódjaihoz tartozó $f_{ij}(\mathbf{P})$, $i=1, 2, \dots, K$, $j=1, 2, \dots, M(i)$ termelési költség függvények lineáris törtfüggvénnyel való közelítését felhasználva az i -edik erőmű j -edik üzemmódjának t -edik periódusbeli, a p_{ijl}^t , $l=1, 2, \dots, r(i, j)$ teljesítmény változók által meghatározott (5.8) szintű üzemelésének költsége:

$$(5.11) \quad a_i \left\{ K_{ij}(x_{i,j-1}^t - x_{ij}^t) + \sum_{l=1}^{r(i,j)} c_{ijl} \cdot p_{ijl}^t \right\}.$$

A teljes tervezési időszakban az erőművi blokkok működéséből származó költséggrész pedig:

$$(5.12) \quad \sum_{t=1}^{27} a_t \sum_{i=1}^K \sum_{j=1}^{M(i)} \left\{ K_{ij}(x_{i,j-1} - x_{ij}^t) + \sum_{l=1}^{r(i,j)} c_{ijl} p_{ijl}^t \right\},$$

ami lineáris függvénye a modell változóinak.

5.2.2. A berendezések állásából, ill. újraindításából származó költséggrész

A 4.4. bekezdésben leírtak szerint az erőművi egységek, ill. az üzemmódok állásából származó költség közelítő értékét a d_{ij}^t mennyiségek állásidőre való összegzésével kapjuk.

Az üzemmód változók definíciója és az üzemmódok megadására tett feltevésünk szerint $x_{ij}^t = 1$ teljesülése esetén az i -edik erőműben, a t -edik periódusban a j -edik üzemmódot követő valamely üzemmód üzemel. Ez azonban azt jelenti, hogy a j -edik üzemmódhoz tartozó legalább egy erőművi egység áll.

A d_{ij}^t mennyiségek állásidőre való összegzése ezek szerint megfelel a $d_{ij}^t x_{ij}^t$ szorzatok teljes tervezési időszakra való összegzésének (kivéve a stagnálási szakaszok második, harmadik és negyedik periódusát).

Az állásköltség tehát:

$$(5.13) \quad \sum_{t=1}^{27} \sum_{i=1}^K \sum_{j=1}^{M(i)} d_{ij}^t x_{ij}^t,$$

ebben az összegzésben $t \neq t_0 + 1, t_0 + 2, t_0 + 3$ és $t \neq t_1 + 1, t_1 + 2, t_1 + 3$.

Ez az összeg azonban tartalmaz az előző tervezési időszak végén már álló egységekhez tartozó költséget is, amittől az egyszerűsítő feltevés szerint eltekintünk. Ennek megfelelően módosítanunk kell az (5.13) értékét. Ehhez változtassuk meg a d_{ij}^t együtthatók definícióját. Legyen $d_{ij}^t = 0$ minden, az első kikapcsolási szakaszhoz, az első bekapcsolási szakaszhoz és az első stagnálási szakasz első periódusához tartozó t érték és minden olyan i, j párral jelzett üzemmód ($i = 1, 2, \dots, K; j = 1, 2, \dots, M(i)$) esetén, amelyre $x_{ij}^0 = 1$. ($x_{ij}^0 = 1$ azt jelenti, hogy az i -edik erőműben az előző tervezési időszak befejezésekor valamely, a j -edik üzemmódot követő üzemmód üzemel, tehát a j -edik üzemmódhoz tartozó valamely erőművi egység áll.)

A d_{ij}^t együtthatók módosított értékével számolva az (5.13) kifejezésben már nem tartozik állásköltség a tervezési időszak első felében — a 10. ábra jelöléseit használva — az ① és ② típusú kikapcsolt egységekhez.

Ugyanakkor a ② és ④ típusú kikapcsolt egységekhez a nap második felében még számolunk állásköltséget, holott az egyszerűsítő feltevés szerint ettől eltekintünk. A ② és ④ típusú kikapcsolás-újraindítást az jellemzi, hogy a kikapcsolt egység az első bekapcsolási szakasz utolsó periódusában sem működik. Jelölje t^* az említett periódus sorszámát.

$x_{i_0 j_0}^{t^*} = 1$ esetén az i_0 -adik erőmű j_0 -adik üzemmódjához ② vagy ④ típusú kikapcsolt egység tartozik. A $d_{i_0 j_0}^t$ értékek teljes tervezési időszakra vonatkozó összegezésakor az (5.13) összegben a nap második felére vonatkozóan az $s_{i_0, j_0}(4 + 2l_0)$ köze-

lító értéke szerepel, feltéve, hogy l_0 a második kikapcsolási szakasz periódusainak a száma. Csökkentsük az (5.13) állásköltséget a

$$(5.14) \quad \sum_{i=1}^K \sum_{j=1}^{M(i)} s_{ij} (4 + 2l_0) x_{ij}^t$$

összeggel. Ha $t=t^*$ esetére a d_{ij}^t együtthatókat módosítjuk úgy, hogy csökkentjük értéküket $s_{ij}(4+2l_0)$ -val, $(i=1, 2, \dots, K; j=1, 2, \dots, M(i))$, akkor (5.13) az állásköltség egyszerűsítő feltevéseink szerinti közelítése.

5.2.3. A hálózati veszteség költsége

Legyen a tekintett feszültségellenőrző periódus a t -edik, ahol nem okoz félreértést, elhagyjuk a periódust azonosító t indexet a változóknál. Jelölje $F(\mathbf{u}, \mathbf{p}, \mathbf{x})$ a hálózati energiaveszteség forint értékét a periódusra, akkor (4.29) alapján a célfüggvény veszteségi része:

$$(5.15) \quad F(\mathbf{u}, \mathbf{p}, \mathbf{x}) = \gamma a_t \left[\mathbf{d}^T \mathbf{u} + \mathbf{b}^T \mathbf{Z}^N \left(\mathbf{D}^* \begin{pmatrix} \mathbf{p} \\ \mathbf{x} \end{pmatrix} + \mathbf{P}^{\min} - \mathbf{P}^F - \mathbf{P}^{*K} \right) \right],$$

ahol γ egy MWh veszteség költsége forintban, a_t a t -edik periódus hossza órában. Az összes hálózati veszteség költség az egyes periódusokra vonatkozó veszteségek költségeinek összege.

5.3. Korlátozó feltételek

Az egyszerűsített modell feltételei is csoportosíthatók aszerint, hogy periódusonként ismétlődően csak egy-egy periódushoz tartozó változók közötti kapcsolatot fejeznek ki, vagy több periódus változói közötti összefüggést írnak elő. A periódusonként ismétlődő feltételek között vannak olyanok, amelyek teljesülését csak az úgynevezett feszültségellenőrző periódusokban követeljük meg explicit módon.

5.3.1. Normál periódus feltételrendszere

A következőkben ismertetésre kerülő feltételek teljesülését minden periódusban meg kell követelnünk. „Normál periódus” azoknak a periódusoknak a neve, amelyekben ezeken a feltételeken kívül további, a periódushoz tartozó változók közötti kapcsolatot biztosító korlátozó feltevés már nem szükséges.

A teljesítmény változók definíciója szerint minden periódusban meg kell követelnünk a

$$(5.16) \quad 0 \leq p_{ijl}^t \leq (P_{ijl}^{\max} - P_{ijl}^{\min})(x_{ij-1}^t - x_{ij}^t)$$

$$i = 1, 2, \dots, K, \quad j = 1, 2, \dots, M(i), \quad l = 1, 2, \dots, r(i, j)$$

összekapcsolási feltételek teljesülését.

Megjegyezzük, hogy annak biztosítására, hogy $p_{ijk}^t > 0$ csak $p_{ijl}^t = P_{ijl}^{\max} - P_{ijl}^{\min}$, $l=1, 2, \dots, k-1$ esetén teljesülhessen, nem kell a modellben korlátozó feltételt ad-

nunk. A termelési költségeket megadó $f_{ij}(P)$ függvények közelítő függvényeinek (4.2) tulajdonsága miatt a modellnek megfelelő vegyesváltozós feladat költségminimumot megvalósító megoldásaira ez automatikusan teljesül.

Az *ellátási feltétel* az egyszerűsített modell változóinak felhasználásával a következőképpen adható meg:

$$(5.17) \quad \sum_{i=1}^K \sum_{j=1}^{M(i)} (x_{ij-1}^t - x_{ij}^t) P_{ij}^{\min} + \sum_{l=1}^{r(i,j)} p_{ijl}^t = P^{t \text{ igény}} + P^{t \text{ veszt}} + P^{t \text{ önf}},$$

Bár erre a tényre az egyszerűsítő feltételek ismertetésénél nem tértünk ki, az erőművi önfogyasztás függvényeket (l. 3.3.1. bekezdés) ebben a modellben az üzemmódoktól függetlennek tételezzük fel s az erőművi teljesítmény lineáris függvényével közelítjük. Így

$$(5.18) \quad P_i^{\text{önf}} = \sum_{i=1}^K P_i^{\text{önf}} \left(\sum_{j=1}^{M(i)} (x_{ij-1}^t - x_{ij}^t) P_{ij}^{\min} + \sum_{l=1}^{r(i,j)} p_{ijl}^t \right),$$

ahol $P_i^{\text{önf}}(P)$ az i -edik erőmű lineáris önfogyasztás függvénye.

Az általános modellben szereplő SOS feltételekre az egyszerűsített modellben nincs szükség. Szükség van azonban az üzemmód definíciója miatt az

$$(5.19) \quad x_{ij-1}^t - x_{ij}^t \geq 0, \quad j = 1, 2, \dots, M(i) - 1,$$

összefüggés megkövetelésére, $i=1, 2, \dots, K$ esetén. Ezekre az *üzemmód változók korlátozó feltételei* néven hivatkozunk.

5.3.2. Feszültségellenőrző periódusok feltételrendszere

A normál periódusban megkövetelt feltételeken kívül a feszültségellenőrző periódusokban további feltételek teljesülését kell megköválnunk.

E helyen a 3.3.1., ill. a 4.5.—4.8. bekezdésekben tárgyaltakra támaszkodva megadjuk az egyszerűsített modellnek az átviteli hálózathoz kapcsolódó feltételrendszerét. Mivel rögzített periódust tekintünk, a felső t azonosító indexet elhagyjuk. A munkapontra vonatkozó mennyiségeket felső $*$ -gal különböztetjük meg.

Feszültségtartási feltételek

Csomópontként adottak a csomóponti feszültségek abszolút értékére vonatkozó $V_i^{\min}, V_i^{\max}, i=1, \dots, N$ korlátok. Ebből a valós részre vonatkozó korlátokat a következőképp számoljuk:

$$(5.20) \quad u_i^{\min} = [(V_i^{\min})^2 - (w_i^*)^2]^{\frac{1}{2}}; \quad u_i^{\max} = [(V_i^{\max})^2 - (w_i^*)^2]^{\frac{1}{2}}; \quad i = 1, \dots, N.$$

A feszültségtartási feltételek a meddőforrás csomópontokra vonatkozóan egyedi alsó-felső korlátok:

$$(5.21) \quad u_i^{\min} \leq u_i \leq u_i^{\max}, \quad i = 1, \dots, L.$$

A fogyasztói csomópontok esetében felhasználhatjuk a (4.15) összefüggést, így a következő feltételrendszer adódik:

$$(5.22) \quad \mathbf{u}^{\min F} \leq \mathbf{v}^{*F} - \mathbf{Y}_4^{-1} \mathbf{Y}_3(\mathbf{u} - \mathbf{v}^{*M}) \leq \mathbf{u}^{\max F},$$

ahol alkalmaztuk 4.6. jelöléseit.

Ágterhelési feltételek

Jelölje \mathbf{T}^{\max} a (3.22)-ben szereplő T_{ik}^{\max} mennyiségek 4.8. bekezdés szerinti, korrigált M -dimenziós vektorát. A (4.23) és (4.18) összefüggések alapján a következő adódik:

$$(5.23) \quad \mathbf{T} \sim \mathbf{BVZ}^N(\mathbf{P}^G - \mathbf{P}^F - \mathbf{P}^{*K}).$$

Figyelembe vesszük a \mathbf{P}^G -re az 5.1.3. bekezdésben kapott összefüggést, így az ágterhelési feltételrendszer a következő lesz:

$$(5.24) \quad -\mathbf{T}^{\max} \leq \mathbf{BVZ}^N \left[\mathbf{D}^* \begin{pmatrix} \mathbf{p} \\ \mathbf{x} \end{pmatrix} + \mathbf{P}^{\min} - \mathbf{P}^F - \mathbf{P}^{*K} \right] \leq \mathbf{T}^{\max}.$$

Megjegyzés: A modellbe mind a fogyasztói csomópontokra vonatkozó feszültségtartási, mind pedig az ágterhelési feltételekből a 4.11. bekezdésben tárgyaltak szerint K_v , ill. K_A számú kerül be.

Meddőforrás feltételek

E feltételrendszer (4.26) és (3.24) alapján azonnal felírható:

$$(5.25) \quad \mathbf{Q}^{\min}(\mathbf{x}) \leq \mathbf{g}^M(\mathbf{v}^*, \mathbf{w}^*) + (\mathbf{Y}_1 - \mathbf{Y}_2 \mathbf{Y}_4^{-1} \mathbf{Y}_3)(\mathbf{u} - \mathbf{v}^{*M}) \leq \mathbf{Q}^{\max}(\mathbf{x}).$$

Itt a 4.8. bekezdés jelöléseit használtuk, és bevezettük a $Q_i^{\min}(\mathbf{x})$, ill. $Q_i^{\max}(\mathbf{x})$, $i = 1, \dots, L$ komponensekből álló, $\mathbf{Q}^{\min}(\mathbf{x})$, ill. $\mathbf{Q}^{\max}(\mathbf{x})$ vektorokat. Meg kell még adnunk az (5.25) jobb és bal oldalán szereplő függvényeket. Az általános modell (3.24) meddőforrás-feltételrendszerével kapcsolatos megfontolások itt is érvényesek, így az \mathbf{x} -től való függés megadásához elegendő tekinteni az általános modell \mathbf{y} üzemiállapot vektora és az \mathbf{x} üzemiállapot vektor közötti lineáris (5.2) összefüggést.

Ennek alapján a $Q_i^{\min}(\mathbf{x})$, $Q_i^{\max}(\mathbf{x})$ függvények a (3.24)-nél alkalmazott jelölésekkel a következő alakot öltik:

$$(5.26) \quad \begin{aligned} Q_i^{\min}(\mathbf{x}) &= (\mathbf{H}\mathbf{x} + \hat{\mathbf{h}})^T \hat{Q}_i^{\min}, & Q_i^{\max}(\mathbf{x}) &= (\mathbf{H}\mathbf{x} + \hat{\mathbf{h}})^T \hat{Q}_i^{\max}, & i \in I_M \cap I_E, \\ Q_i^{\min}(\mathbf{x}) &= Q_i^{\min}, & Q_i^{\max}(\mathbf{x}) &= Q_i^{\max}, & i \in I_M \setminus I_E. \end{aligned}$$

5.3.3. A periódusokat összekapcsoló feltételek

Az egyszerűsített modellben nincs szükség az általános modell kikapcsolási feltételeinek megfelelő feltételekre. Elegendő ebben a pontban a 4.1. bekezdésben leírt üzemeltetési rendszabályok matematikai megfogalmazását megadnunk.

A kikapcsolási szakaszban nem engedünk meg bekapcsolást. Ha a t -edik periódus a bekapcsolási szakasz utolsó periódusa, vagy kikapcsolási szakaszhoz tartozik, akkor

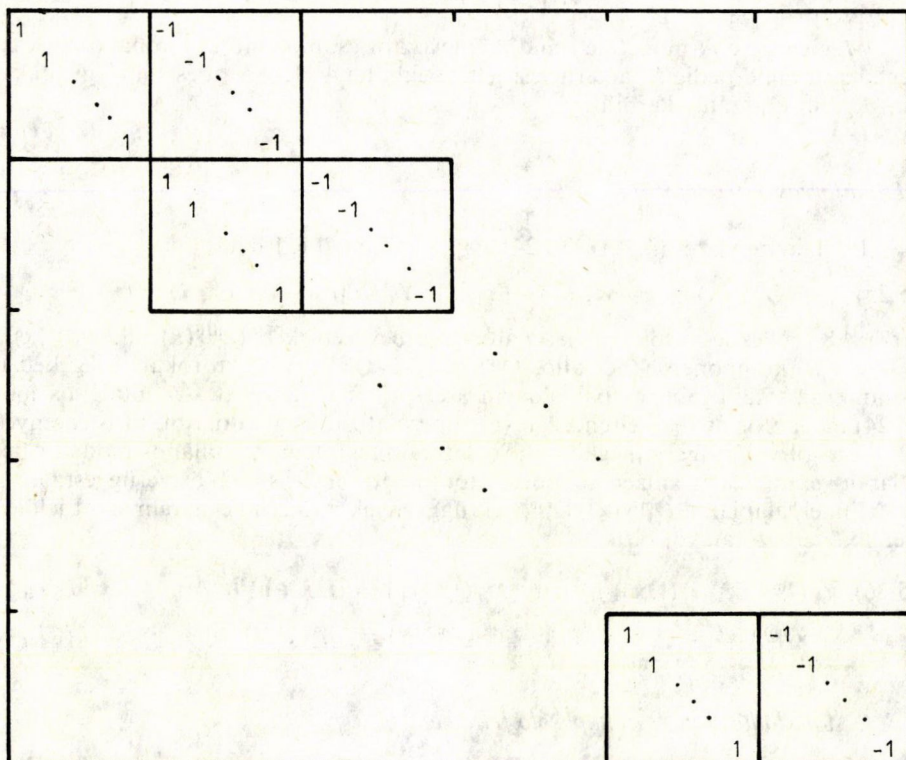
$$x_{ij}^t = 1 \quad \text{esetén} \quad x_{ij}^{t+1} = 1$$

szükségképpen. Ezt a logikai feltételt az

$$(5.27) \quad x_{ij}^t \leq x_{ij}^{t+1}$$

egyenlőtlenség írja le a legegyszerűbben. A feltételek mátrixát a 13. ábra mutatja.

a megelőző bekapcsolási szakasz utolsó periódusának üzemmod változói	a kikapcsolási szakasz 1. periódusá- nak üzemmod változói	a kikapcsolás szakasz 2. periódusá- nak üzemmod változói	...	a kikapcsolási szakasz utolsó periódusának üzemmod változói	a következő stagnálási szakasz első periódusának üzemmod változói
---	---	--	-----	---	--



13. ábra

A kikapcsolási feltételek mátrixa

5.3.4. Tüzelőanyagkényszer feltételek

A tüzelőanyagkényszer feltételek is periódusokat összekapcsoló feltételek. Ismeretükhöz elegendő az általános modellnél a 3.3.2. bekezdésben leírtakra hivatkozni és a (3.32) formulát az egyszerűsített modell változóival felírni:

$$(5.29) \quad R_{i \min} \leq \sum_{t=1}^{27} a_t \sum_{j=1}^{M(i)} \left\{ p_{ij}^{\min} (x_{ij-1}^t - x_{ij}^t) + \sum_{l=1}^{r(i,j)} p_{ijl}^t \right\} \leq R_{i \max}.$$

Legyen

$$(5.30) \quad h_i(\mathbf{p}, \mathbf{x}) = \sum_{t=1}^{27} a_t \sum_{j=1}^{M(i)} \left\{ p_{ij}^{\min} (x_{ij-1}^t - x_{ij}^t) + \sum_{l=1}^{r(i,j)} p_{ijl}^t \right\},$$

a \mathbf{p} teljesítmény változó és \mathbf{x} üzemmód változó lineáris függvénye. Ezt felhasználva a tüzelőanyagkényszer feltételek alakja:

$$(5.31) \quad R_{i \min} \leq h_i(\mathbf{p}, \mathbf{x}) \leq R_{i \max}.$$

5.4. Az egyszerűsített modell szerkezete, jellemzői, mérete

Az egyszerűsített modellnek megfelelő nagyméretű, vegyesváltozós, lineáris cél-függvényt és feltételeket tartalmazó matematikai programozási feladat a következő.

A minimalizálandó célfüggvényt az (5.13)-ban szereplő állásköltség (5.14)-gyel csökkentett értékének és az (5.12) termelési költségnek, valamint az (5.15) hálózati veszteség költségnek az összege szolgáltatja.

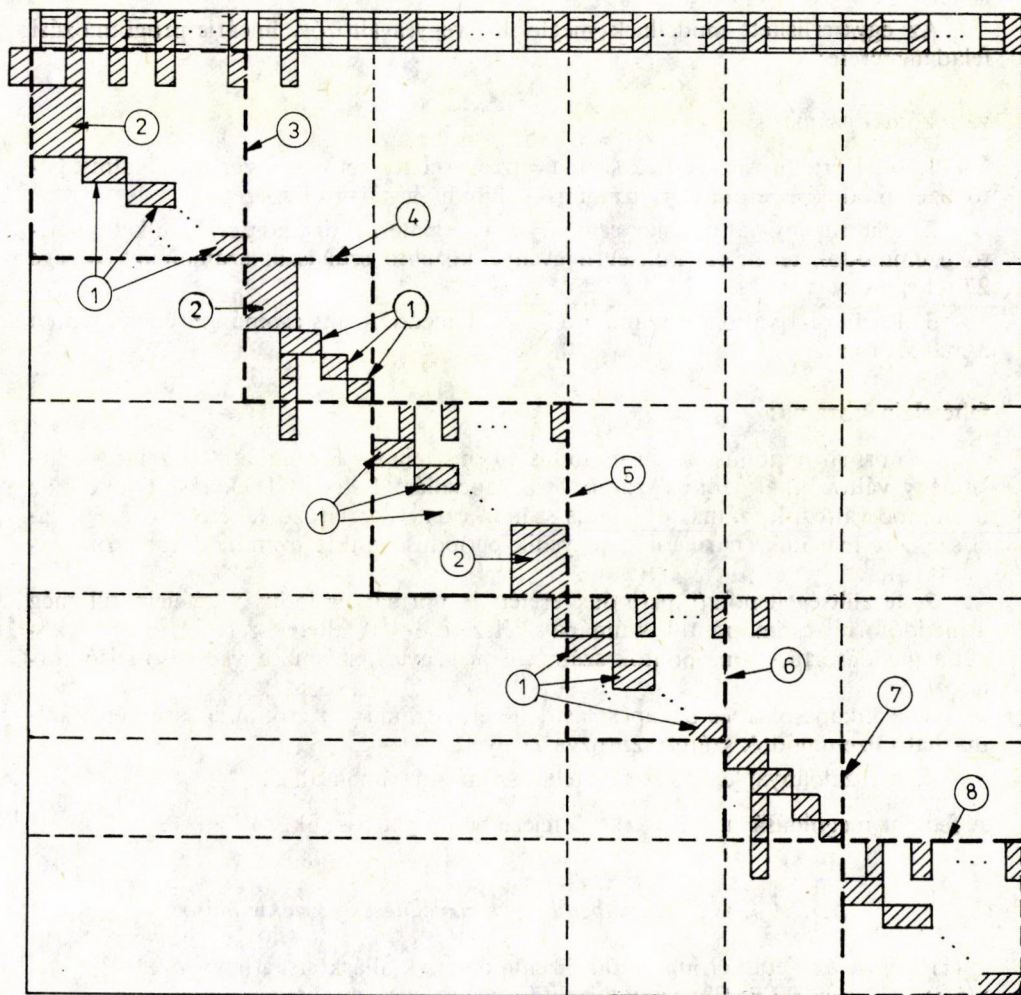
A korlátozó feltételrendszer a következőket tartalmazza:

- (5.16) összekapcsolási feltételek,
- (5.17) ellátási feltétel,
- (5.19) üzemmód változók korlátozó feltételei,
- (5.22) feszültségtartási feltételek,
- (5.24) ágterhelési feltételek,
- (5.25) meddőforrás feltételek,
- (5.27)–(5.28) kikapcsolási-bekapcsolási feltételek,
- (5.30) tüzelőanyagkényszer feltételek.

A 15. ábrán e feladat együtthatómátrixának rajza látható. A változók periódusonként a következő sorrendben szerepelnek: feszültségi változók, teljesítmény változók, üzemmód változók. (Emlékeztetőül megemlítjük, hogy csak néhány periódushoz tartoznak feszültségi változók és a stagnálási szakaszok második, harmadik és negyedik periódusaihoz nem tartoznak üzemmód változók.)

Minden periódushoz tartozik a feltételeknek egy blokkja. Ebben azok a feltételek szerepelnek, amelyek csak az adott periódus változóit tartalmazzák. Normál periódusokhoz tartozó blokkot az ábrán ① jelöli, míg a feszültségellenőrző periódusok blokkját ②.

Az ábrán szaggatott vonallal egybekapcsolt blokkok egy-egy szakasz (kikapcsolás, stagnálás, bekapcsolás) feltételeit foglalják össze. Az ① és ②-vel jelölt blokkokon kívül a kikapcsolási szakasznak megfelelő blokkok (ezeket az ábrán ③ és ⑥ jelöli) a kikapcsolási feltételeket, a bekapcsolási szakaszoknak megfelelő blokkok (⑤, ⑧)



15. ábra

Az egyszerűsített modell szerkezete

a bekapcsolási feltételeket tartalmazzák. A stagnálási szakaszoknak megfelelő blokkok (jelük az ábrán ④ és ⑦) egyetlen üzemmódváltó vektort tartalmaznak. A hozzájuk tartozó, a periódusoknak megfelelő „kis blokkok” ezen üzemmódváltó vektoron keresztül kapcsolódnak egymáshoz.

Ugyanezek az üzemmódváltó vektorok létesítenek kapcsolatot a ③ és ⑤ blokk, ill. a ⑥ és ⑧ blokk között. Az ⑤ és ⑥ blokk az első bekapcsolási szakasz utolsó periódusához tartozó üzemmódváltó vektoron keresztül kapcsolódik egymáshoz. Az ábra felső részén levő, az összes periódust összekapcsoló feltételek a tüzelőanyag-kényszer feltételek.

Az egyszerűsített modellnek megfelelő vegyesváltozós, lineáris programozási feladat mérete:

változóinak száma:

1. 0–1 értékű változóinak száma = az energiatermelő rendszerben alkalmazható üzemmódok és erőművek számának különbsége 21-gyel szorozva;

2. teljesítmény változóinak száma = az energiatermelő rendszerben alkalmazható üzemmódok termelési költségfüggvényei-közelítő szakaszai számainak összege 27-tel szorozva;

3. feszültségi változók száma = a hálózat meddőforrás-csomópontjainak száma 3-mal szorozva;

feltételeinek száma:

1. normál periódus esetén periódusonként (tehát a feladatban 24-szer): a teljesítmény változókkal azonos számú összekapcsolási feltétel + 1 ellátási feltétel + az üzemmód változók számával azonos számú, ezeket korlátozó feltétel. (A stagnálási szakasz második, harmadik, negyedik periódusában az üzemmód változók száma 0.);

2. feszültségellenőrző periódus esetén a normál periódus feltételein túl még a meddőforrás-csomópontok számának kétszerese + 60 feltétel szerepel, amelyek közül a meddőforrás-csomópontok számával megegyező számú, egyedi alsó-felső korlát;

3. a kikapcsolási és bekapcsolási feltételek száma = az erőműrendszerben alkalmazható üzemmódok száma szorozva 21-gyel;

4. a tüzelőanyagkényszer feltételek száma maximálisan 5.

A feladat megoldásával a következő fejezetben foglalkozunk.

5.5. A 4. és 5. fejezetben bevezetett jelölések összefoglalása

$s_{ij}(\tau)$	az i -edik erőmű j -edik üzemmódjának állásköltségfüggvénye,
$r(i, j)$	az $f_{ij}(P)$ függvény közelítő szakaszainak a száma,
$P_{ijl}^{\min}, P_{ijl}^{\max}$	az l -edik közelítő szakasz végpontjaihoz tartozó teljesítményszintek,
c_{ijl}	az l -edik közelítő szakasz meredeksége,
d_{ij}^t	az i -edik erőmű j -edik üzemmódja állásköltségének a t -edik periódusban figyelembe veendő része,
x^t, x_{ij}^t	a t -edik periódushoz tartozó üzemmód változó, ill. komponens,
x^0	a tervidőszakot megelőző nap utolsó periódusának üzemmódváltozó vektora,
t_0, t_1	az első és a második stagnálási szakasz kezdetének megfelelő periódus sorszáma,
p_{ijl}^t	$l = 1, 2, \dots, r(i, j)$ az i -edik erőmű j -edik üzemmódja teljesítményszintjét megadó teljesítmény változó komponense — a t -edik periódusban,
p^t, p	a t -edik periódushoz tartozó teljesítményváltozó vektor és a teljes időszakhoz tartozó teljesítményváltozó vektor,

t^*	az első bekapcsolási szakasz utolsó periódusának sorszám,
$h_i(\mathbf{p}, \mathbf{x})$	az i -edik erőmű napi termelését megadó lineáris függvény,
L	a meddőforrás csomópontok száma ($=N_M$),
N_F	a fogyasztói csomópontok száma, $L + N_F = N$,
\mathbf{u}	a feszültségi változók vektora valamely rögzített periódusban, komponenseit a meddőforrás-csomópontok feszültségeinek valós része alkotja, $\mathbf{u} \in R^L$,
$(\mathbf{v}^*, \mathbf{w}^*)$	a munkaponti feszültségek vektora,
u_i^{\min}	a meddőforrás-csomópont feszültsége valós részének alsó korlátja,
u_i^{\max}	a meddőforrás-csomópont feszültsége valós részének felső korlátja,
\mathbf{B}	$M \times M$ méretű diagonális mátrix, diagonális elemei a B_{ik} mennyiségek (lásd F3. rész),
\mathbf{V}	$N \times N$ méretű mátrix, a teljesítményáramlás feszültség-függésének kompakt felírása céljából a (4.17) összefüggés után bevezetve (definícióját lásd ott),
\mathbf{A}	a hálózat — a Függelék jelöléseivel az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ gráf-csúcs, él — incidencia mátrixa,
\mathbf{Z}^N	a következőképpen kapható: törlendő az $\mathbf{A}^T \mathbf{B} \mathbf{V}$ mátrix referenciaponti sora és oszlopa, az így adódó mátrix inverze $\mathbf{0}$ sorral és oszloppal egészítendő ki a törölt soroknak megfelelő pozíciókban,
$\mathbf{Y} = \left(\begin{matrix} \mathbf{Y}_1 & \mathbf{Y}_2 \\ \mathbf{Y}_3 & \mathbf{Y}_4 \end{matrix} \right)_{\substack{L \\ N_F}}^L$	a $\mathbf{g}(\mathbf{v}, \mathbf{w}^*): R^N \rightarrow R^N$ leképezés <i>Jacobi-mátrixa</i> , a meddőforrások és fogyasztói csomópontok szerint particionálva,
\mathbf{D}^*	a teljesítmény és üzemállapot változókból a tényleges csomóponti hatásosteljesítmény-betáplálást adó lineáris transzformáció mátrixa,
\mathbf{P}^K	az ágankénti hatásosteljesítmény-veszteségek korrekciós vektora,
\mathbf{d}^v	a célfüggvény veszteségi részében \mathbf{u} együtthatója,
\mathbf{b}^v	a célfüggvény veszteségi részében a \mathbf{p}, \mathbf{x} változók együtthatóinak konstrukciójához szükséges vektor,
γ	egy MWh hatásosteljesítmény-veszteség költsége Ft-ban.

6. A napi menetrend meghatározása

A villamos energiát termelő erőműrendszer és az azt továbbító hálózat napról napra változik, így mind a megoldandó feladat, mind a feladat mérete más és más lesz, bár az egyes feladatok felépítése hasonló. Ebben a fejezetben azt ismertetjük, hogy az ütemezési probléma egyszerűsített modelljének megfelelő napi feladatot hogyan lehet előállítani és megoldani.

6.1. A napi adatoknak megfelelő vegyesváltozós feladat generálása

A hálózatban szükséges javítások, az üzemzavarok elhárítása, vagy más szerelési munkák miatt egy adott napon a hálózat egyes ágainak kapacitása megváltozik (pl. párhuzamos vezetékek, transzformátorok esetében), vagy pedig egyáltalán nem is használható villamos energia továbbítására. Hasonló a helyzet az erőművekkel; az

üzemeltethető erőművek, ill. a használható üzemmódok is napról napra változnak nagyjavítások, berendezések elromlása stb. miatt. Ezért az ütemezési problémának megfelelő napi feladatot mindig elő kell állítani, vagyis a villamos hálózat és az erőművek állandó adatait tartalmazó adatbázisból a megfelelő módosítások felhasználásával meg kell határozni a naprakész állapotot, az adott napon használható villamos hálózatot, az egyes ágak aktuális kapacitását, az erőművek adott napon lehetséges üzemmódjait és az azokhoz tartozó korlátokat valamint a feltételeket.

Külön problémát képez egy adott napi feladatban fellépő feszültségellenőrző periódusok feltételrendszerének előállítása. Ehhez először megoldunk egy *load flow feladatot* a munkapont és az ehhez tartozó hálózati veszteségek meghatározása céljából, majd az ezen munkaponthoz és veszteséghez tartozó feltételrendszert generáljuk (a feladatban a feltételek közül csak bizonyosakat veszünk figyelembe, lásd 4.11. bekezdést). Megjegyezzük, hogy az optimalizálás eredményeként a meddőforrást tartalmazó csomópontokban jelentkező feszültségértékek alapján újra kiszámítjuk a feszültség-eloszlást. Mivel a célfüggvényben szerepelnek veszteségi együtthatók, ezért a munkaponti veszteségeknél kisebbeket kapunk (lásd 4.5. bekezdést).

A napi feladat többi részét, a ki- és bekapcsolási feltételeket, az összekapcsolási, az ellátási és a tüzelőanyag-kényszer feltételeket a már említett adatokból, ill. az adott napra előrebecsült fogyasztói villamosenergia-igényekből határozzuk meg. Itt jegyezzük meg, hogy az előrebecslés legfeljebb 1–2%-kal tér el a tényleges igénytől.

6.2. Megoldási lehetőségek

Az ilyen módon előállított nagyméretű, vegyesváltozós — mind valós, mind csak egész (0–1) értékeket felvevő változókat tartalmazó — modell számítógépes optimalizálását többféle algoritmussal lehet elvégezni.

Előszörban a *Benders dekompozíciós eljárás* felhasználására gondoltunk, amellyel az egész feladatot egyszerre oldottuk volna meg. Ezt az elgondolást némi megfontolás után elvetettük. A *Benders dekompozíció* folyamán ugyanis (lásd 6.4. bekezdés) nagyszámú, egészértékű változót tartalmazó feladatot kell minden iterációban megoldani, amelyekből pedig már egy feladatnak a megoldása önmagában is nehéz feladat lett volna. Az ilyen, csak egészértékű változókat tartalmazó feladatoknak már nincsen meg az a dekomponálható struktúrája sem, amely az eredeti feladatra — néhány feltétel elhagyása után — jellemző. Szóba került a feladat *branch and bound* (korlátozás és szétválasztás) módszerrel történő megoldása is, amelyben a keletkezett lineáris programozási feladatokat a *Dantzig—Wolfe dekompozíciós eljárással* optimalizáljuk. Ezt a lehetőséget azért vetettük el, mert az egészértékű változók száma igen nagy lehet (esetenként 400 is), ami pedig a *branch and bound eljárást* teszi gyakorlatilag felhasználhatatlanná.

Mindezen megfontolásokra és az ütemezési probléma fizikai hátterére támaszkodva, egy részben heurisztikus elemeket is tartalmazó, dekompozíciós, optimalizáló eljárás alkalmazása mellett döntöttünk: a feladatot durván szólva periódusonként oldjuk meg, de megfelelő mellékfeltevésekkel biztosítjuk ezek összekapcsolhatóságát és a tüzelőanyagkényszer feltétel teljesülését. Az eljárást a következő pontban részletezzük.

6.3. Az optimalizáló eljárás

Az ütemezési probléma egyszerűsített modelljének megfelelő feladat feltételeit nem explicit formában — minden feltételt egyszerre — vesszük figyelembe, hanem csoportokra bontva. A szétbontást az teszi lehetővé, hogy a ki- és bekapcsolási feltételektől eltekintve az egyes periódusokat csak a tüzelőanyagkényszer feltételek fogják össze. Az optimalizáló algoritmus a következő lépésekből áll (az eljárást arra az esetre részletezzük, amikor egyetlen tüzelőanyagkényszerszerű előíró feltételünk van).

1. Elhagyjuk a tüzelőanyagkényszer feltételét.

2. A fennmaradó nagyméretű, vegyes, egészértékű programozási feladatot — amelyben az egyes periódusok közötti kapcsolatot a ki- és bekapcsolási feltételek, valamint a stagnálási szakaszok üzemmód változói biztosítják — a következő módon oldjuk meg (az algoritmus 3. és 4. lépése).

3. Megoldjuk rendre az első, második és harmadik feszültségellenőrző periódust úgy, hogy az üzemmód változók lehetséges értékeire nézve bizonyos korlátozásokkal élünk. Az első feszültségellenőrző periódus megoldásában minden, az adott napon alkalmazható üzemmód előfordulhat, nincsen korlátozás. A második feszültségellenőrző periódus (amely az első stagnálás első periódusa) megoldásában csak olyan üzemmódokat engedünk meg, amelyek az első feszültséges periódus megoldásában kapott üzemmódokból kikapcsolásokkal kapható meg. Végül a harmadik feszültségellenőrző periódusban (a délelőtti legnagyobb teljesítményigény periódusában) olyan üzemmódok fordulhatnak elő, amelyek a második feszültségellenőrző periódus megoldásában szereplő üzemmódokból bekapcsolásokkal nyerhetők.

4. Ezek után megoldjuk sorban a még meg nem oldott periódusok feladatát úgy, hogy az adott periódus előtt, ill. után levő, már megoldott periódusok üzemmód változóinak értékeit is figyelembe vesszük. Ez úgy történik, hogy kikapcsolási szakaszban csak olyan üzemmódokat vesszünk figyelembe az adott periódus megoldásánál, amelyek (i) a korábbi, már megoldott és a kikapcsolási szakaszban levő periódus üzemállapot-rendszeréből kikapcsolással jöhetnek létre (ezt röviden úgy nevezzük, hogy az előző periódusból kikapcsolással kaphatók meg), és amelyekből (ii) a későbbi, már megoldott és a kikapcsolási szakaszban levő periódus üzemállapot rendszere további kikapcsolással kapható meg. Hasonlóképpen bekapcsolási szakaszban csak olyan üzemmódokat engedünk meg egy adott periódus megoldásánál, amelyek (i) a korábbi, már megoldott és a bekapcsolási szakaszban levő periódus üzemállapot rendszeréből bekapcsolással jöhetnek létre (ezt röviden úgy nevezzük, hogy egy előző periódusból bekapcsolással kapható meg) és amelyekből (ii) a későbbi, már megoldott és a bekapcsolási szakaszban levő periódus üzemállapot rendszere további bekapcsolással kapható meg. Mind a kikapcsolásnál, mind a bekapcsolásnál lehetnek (és vannak is) nem változó üzemmód változók.

Például a második periódus megoldásánál csak olyan üzemmódok jöhetnek számításba, amelyek az első feszültségellenőrző periódusból kikapcsolással jöhetnek létre, és amelyekből a második (már szintén megoldott) feszültségellenőrző periódus üzemmódjai csak kikapcsolásokkal valósíthatók meg. A harmadik periódus megoldásánál a második periódus és a második feszültségellenőrző periódus üzemmódjait kell figyelembe venni stb. A harmadik feszültségellenőrző periódus után de még a má-

sodik stagnálás előtti periódusok megoldásánál olyan üzemmódokat engedünk meg, amelyek csak az előzőleg megoldott periódusoktól függenek, abból kikapcsolással állíthatók elő, míg a stagnálási szakasz utáni periódusok megoldásánál olyan üzemmódokat engedünk meg, amelyek az előzőleg megoldott periódusból bekapcsolással állíthatók elő.

Az egy periódusnak megfelelő feladatot mindig a következő, 6.4. bekezdésben leírásra kerülő *Benders dekompozíciós eljárással* oldjuk meg, amely képes a fentebb leírt üzemmód korlátozásokat figyelembe venni.

5. Az egész napra vonatkozó megoldás fentebbiekben leírt meghatározása után ellenőrizzük, hogy a megoldás kielégíti-e a tüzelőanyagkényszer feltételt. Ha igen, akkor az algoritmus véget ér, megkaptuk az optimális megoldást, egyébként pedig az algoritmus 6. lépésben leírásra kerülő iteratív eljárást alkalmazzuk a kapott megoldás módosítására.

6. Az iterációs eljárásban lényegében a tüzelőanyag költség növelésével, ill. csökkentésével érjük el a tüzelőanyagkényszeres erőmű villamosenergia-termelésének csökkentését, ill. növelését.

Az 5.3.4. bekezdésben szereplő (5.31) képlet szerint a tüzelőanyagkényszer feltétele a

$$(6.1) \quad R_{i \min} \leq h_i(\mathbf{p}, \mathbf{x}) \leq R_{i \max}$$

alakban adható meg. Itt $(R_{i \min} + R_{i \max})/2$ egy előre megadott érték, továbbá

$$\frac{R_{i \max} + R_{i \min}}{2} - R_{i \min} = R_{i \max} - \frac{R_{i \max} + R_{i \min}}{2} = \frac{k}{100} \frac{R_{i \max} + R_{i \min}}{2},$$

ahol jelen esetben k egy előre megadott egész szám, amely a megengedett eltérést adja meg %-ban (a szokásos érték: $k=3, 4, 5$). Az adatelőkészítés folyamán arról külön meggyőződünk, hogy az adott $R_{i \min}$ és $R_{i \max}$ állandók mellett lehetséges-e olyan (\mathbf{p}, \mathbf{x}) vektorpárt találni, amely a (6.1) egyenlőtlenséget kielégíti.

Jelöljük most az algoritmus 3. és 4. pontjának végrehajtása során meghatározott megoldást $(\mathbf{p}_0, \mathbf{x}_0)$ -lal. Ha most

$$R_{i \min} \leq h_i(\mathbf{p}_0, \mathbf{x}_0) \leq R_{i \max}$$

egyenlőtlenség nem teljesül (egyébként már az algoritmus 5. pontjában végetért volna az optimalizálás), akkor lerögzítjük az \mathbf{x}_0 vektort, vagyis a továbbiakban nem változtatunk üzemmódot. Tekintsük azt a feladatot, amelyet az eredeti feladatból úgy kapunk, hogy a tüzelőanyagkényszer feltételt elhagyjuk, és az \mathbf{x}_0 vektor értékeit behelyettesítjük. Jelöljük ezt a feladatot $F_0(\mathbf{x}_0)$ -lal, ez egy 27, egymástól független blokkból álló közönséges lineáris programozási feladat a \mathbf{p} ismeretlen vektorra nézve. Ebből a feladatból kiindulva képezzük az $F_1(\mathbf{x}_0), F_2(\mathbf{x}_0), \dots$ feladatok sorozatát, az egyes feladatok egymástól csak a tüzelőanyagkényszeres erőmű működésben levő üzemmódjának teljesítmény változójához tartozó célfüggvénybeli c_1, c_2, \dots, c_k együtthatóban különböznek. Az $F_k(\mathbf{x}_0)$ feladatban a c_k együttható szerepel, az $F_0(\mathbf{x}_0)$ feladatban az eredeti feladatban szereplő, c_0 -lal jelölt együttható. Az $F_k(\mathbf{x}_0)$ feladat optimális megoldását \mathbf{p}_k -val jelölve a c_{k+1} együtthatót és ezzel együtt az $F_{k+1}(\mathbf{x}_0)$

feladatot a következő rekurziók adják meg ($s_0=1$):

$$(6.2) \quad c_{k+1} = c_k s_{k+1} / s_k$$

$$(6.3) \quad s_{k+1} = \begin{cases} s_k \left[1 - 0,3 \frac{R_{i\min} - h_i(\mathbf{p}_k, \mathbf{x}_0)}{\frac{R_{i\min} + R_{i\max}}{2}} \right], & \text{ha } h_i(\mathbf{p}_k, \mathbf{x}_0) < R_{i\min} \\ s_k \left[1 - 0,3 \frac{R_{i\max} - h_i(\mathbf{p}_k, \mathbf{x}_0)}{\frac{R_{i\min} + R_{i\max}}{2}} \right], & \text{ha } h_i(\mathbf{p}_k, \mathbf{x}_0) > R_{i\max}. \end{cases}$$

A többszörös munkát elkerülendő már előfordult árakat kizárunk. Ezt azáltal érjük el, hogy két, egyre kisebb ingadozást engedő s_{\min} és s_{\max} korlát között hagyjuk csak az s_k szorzót változni. Legyen kezdetben

$$s_{\min} = 0,001 \quad s_{\max} = 1000.$$

Egy s_k szorzó alultermelést, ill. túltermelést okoz, ha az s_k -val számított c_k együtthatóval felírt modell optimális \mathbf{p}_k megoldása esetén $h_i(\mathbf{p}_k, \mathbf{x}_0) < R_{i\min}$, (ill. $h_i(\mathbf{p}_k, \mathbf{x}_0) > R_{i\max}$). Amennyiben az újonnan meghatározott \mathbf{p}_k vektor esetén alultermelés van, akkor $s_{\min} < s_k < s_{\max}$ esetén $s_{\max} = s_k$ lesz az új határ (túltermelés esetén $s_{\min} = s_k$). Tehát s_{\max} tartalmazza az eddigi alultermelést okozó szorzók közül a legkisebbet, s_{\min} pedig a túltermelést okozó szorzók közül a legnagyobbat. Ha az $s_{\min} \leq s_k \leq s_{\max}$ egyenlőtlenség teljesül a (6.3) által kiszámított s_k együtthatóra, akkor a (6.2) szerint árat módosítunk, egyébként az $s_k = \frac{s_{\min} + s_{\max}}{2}$ értéket használjuk a c_k meghatározá-

sára (6.2)-ben. Ez az s_{\min} és s_{\max} kettős korlát állandóan szűkíti a lehetséges szorzók választási intervallumát, tehát egyszer már használt c_k ár nem térhet vissza.

A leírt ármódosításnak a következő az értelme. Ha a tüzelőanyagkényszeres erőműben kevesebb villamos energiát termelünk, mint kellene (alultermelés van, vagyis $h_i(\mathbf{p}_k, \mathbf{x}_0) < R_{i\min}$), akkor a következő feladatban a célfüggvénybeli együtthatót az előzőhöz képest csökkentjük, hogy a termelt energia növekedjék (ugyanis a modell költség-minimalizáló). Hasonlóképpen túltermelés esetén (vagyis ha $h_i(\mathbf{p}_k, \mathbf{x}_0) > R_{i\max}$) megdrágítjuk a következő, $(k+1)$ -edik feladatban az energia termelésének költségét, hogy az adott erőműben kevesebb energiát termeljünk.

Az $F_0(\mathbf{x}_0)$, $F_1(\mathbf{x}_0)$, $F_2(\mathbf{x}_0)$... feladatok megoldásaként nyert $\mathbf{p}_0, \mathbf{p}_1, \dots$ optimális megoldások sorozatára a következő két eset közül az egyik fennáll:

a) találunk egy olyan \mathbf{p}_k megoldást, amelyre a (6.1) egyenlőtlenség teljesül, azaz a $(\mathbf{p}_k, \mathbf{x}_0)$ megoldás kielégíti a tüzelőanyagkényszer feltételét is;

b) találunk olyan \mathbf{p}_i és \mathbf{p}_j megoldaspárt, amelyre a $h_i(\mathbf{p}_i, \mathbf{x}_0) < R_{i\min}$ és $h_i(\mathbf{p}_j, \mathbf{x}_0) > R_{i\max}$ teljesül.

Egyéb esetek — például csak alultermelést okozó \mathbf{p}_k megoldások előfordulása — az adatoknak a (6.1) képlet megadása után vázolt ellenőrzése miatt nem fordulhatnak elő. Az a) esetben az optimalizálás végetér, a \mathbf{p}_k megoldásvektort, ill. a $(\mathbf{p}_k, \mathbf{x}_0)$ megoldást az eredeti feladat optimális megoldásának tekintjük. A b) esetben a következő-

képpen járunk el. A $h_i(\mathbf{p}, \mathbf{x}_0)$ függvény \mathbf{x}_0 lerögzítése után \mathbf{p} -ben lineáris. Így

$$\alpha h_i(\mathbf{p}_i, \mathbf{x}_0) + (1 - \alpha) h_i(\mathbf{p}_j, \mathbf{x}_0) = \frac{R_{i\min} + R_{i\max}}{2}$$

egyenlőségből az α szorzó $0 \leq \alpha \leq 1$ meghatározható és ezzel a $\mathbf{p}_{\text{opt}} = \alpha \cdot \mathbf{p}_i + (1 - \alpha) \mathbf{p}_j$ vektor is, amelyet optimális megoldásnak tekintünk, hiszen ez megengedett és

$$h_i(\mathbf{p}_{\text{opt}}, \mathbf{x}_0) = \frac{R_{i\min} + R_{i\max}}{2}$$

is fennáll.

Természetesen α helyett használhatjuk a következő egyenlőségekből nyerhető α_1, α_2 számokat is:

$$(6.4) \quad \alpha_1 h_i(\mathbf{p}_i, \mathbf{x}_0) + (1 - \alpha_1) h_i(\mathbf{p}_j, \mathbf{x}_0) = R_{i\min},$$

$$\alpha_2 h_i(\mathbf{p}_i, \mathbf{x}_0) + (1 - \alpha_2) h_i(\mathbf{p}_j, \mathbf{x}_0) = R_{i\max},$$

ahol α_1 és α_2 a tüzelőanyagkényszer feltétel alsó és felső korlátjához tartozó két érték.

Az α szám használatát az indokolja, hogy a tüzelőanyagkényszer feltétel egyenlőség alakjában is megadható egy adott mennyiségű tüzelőanyag elfogyasztásának előírásával, ill. egy adott $(R_{i\min} + R_{i\max})/2$ mennyiségű energia fejlesztésének megkövetelésével.

6.4. A Benders dekompozíció a részfeladatok megoldására

A megelőző 6.3. bekezdésben ismertetett optimalizáló eljárás alkalmazása esetén a teljes feladat 27 részfeladatra bomlik fel. Ezek — típusuk szerint — három csoportba sorolhatók:

- a) feszültségellenőrző periódusok feladata;
- b) stagnálási szakaszok nem első periódusainak feladata;
- c) a megmaradó, ún. normál periódusok feladata.

Az a) és c) pont alatti feladatok lineáris, vegyes, egészértékű optimalizálási feladatok, ezek megoldására a feladatok sajátosságait felhasználó *Benders dekompozíció*t alkalmazunk. A b) pontban említett feladatok egyfeltételes, korlátos változójú lineáris programozási feladatok, megoldásukra a „*mohó*” algoritmust alkalmazzuk (KOVÁCS LÁSZLÓ BÉLA [40]).

A fejezet további részeiben a *Benders dekompozíció*t és annak alkalmazásakor végrehajtott egyszerűsítéseket írjuk le. Leírásunkban csupán a lineáris, vegyes, egészértékű esetre szorítkozunk; a részletek iránt érdeklődő olvasónak BENDERS [1], KOVÁCS LÁSZLÓ B. [40], LASDON [43] művét ajánljuk.

A *Benders dekompozíció* a következő típusú optimalizálási feladat megoldására szolgál (az algoritmus és a feladat leírásakor jelöléseink függetlenek a korábbi fejezet-

ben bevezetett változó-elnevezésektől).

$$(6.5) \quad \begin{aligned} \mathbf{Ax} + \mathbf{Fy} &\leq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \\ \mathbf{y} &\in Y \\ \max (\mathbf{c}^T \mathbf{x} + \mathbf{f}^T \mathbf{y}), \end{aligned}$$

ahol: \mathbf{A} : $m \times n_1$ -es mátrix;
 \mathbf{F} : $m \times n_2$ -es mátrix;
 \mathbf{c}, \mathbf{x} : n_1 dimenziós vektor;
 \mathbf{f}, \mathbf{y} : n_2 dimenziós vektor;
 \mathbf{b} : m dimenziós vektor;
 Y : n_2 dimenziós diszkrét halmaz.

A kitűzött feladat ekvivalens az

$$(6.6) \quad \begin{aligned} x_0 + ((\mathbf{u}^j)^T \mathbf{F} - \mathbf{f}^T) \mathbf{y} &\leq (\mathbf{u}^j)^T \mathbf{b}, \quad j = 1, \dots, p \\ (\mathbf{v}^j)^T \mathbf{Fy} &\leq (\mathbf{v}^j)^T \mathbf{b}, \quad j = 1, \dots, r \\ \mathbf{y} &\in Y \\ \max x_0 \end{aligned}$$

feladattal, ahol $\mathbf{u}^j, j = 1, \dots, p$, ill. $\mathbf{v}^j, j = 1, \dots, r$ az

$$(6.7) \quad \begin{aligned} \mathbf{A}^T \mathbf{u} &\leq \mathbf{c} \\ \mathbf{u} &\geq \mathbf{0} \end{aligned}$$

feltételekkel meghatározott poliéder extrémális pontjai, ill. extrémális irányai.

A (6.5) vegyes változós feladat helyett a (6.6) — az x_0 változótól eltekintve — egészértékű feladatot kívánjuk megoldani. Ez utóbbira relaxációt alkalmazunk a feltételek explicit alakban való előállíthatatlansága és nagy száma miatt. Így egy iteratív eljárást kapunk, amelyben az i -edik iteráció a (6.6) feladat egy — feltételek elhagyásával keletkező — relaxáltjának (optimális megoldása legyen \mathbf{y}^i), valamint az

$$(6.8) \quad \begin{aligned} \mathbf{A}^T \mathbf{u} &\leq \mathbf{c} \\ \mathbf{u} &\geq \mathbf{0} \\ \min (\mathbf{b} - \mathbf{Fy}^i)^T \mathbf{u} \end{aligned}$$


lineáris programozási feladat megoldásából áll.

A (6.8) feladatok megengedett tartománya azonos, az iterációk során csak a cél-függvényvektor változik a (6.6) feladat utolsóként megoldott relaxáltjának \mathbf{y}^i optimális megoldása függvényében. A (6.8) feladatok az optimalitási kritérium teljesülésének ellenőrzésére és a (6.6) feladat újabb és újabb feltételének előállítására szolgálnak.

Az iteratív eljárás végén a (6.5) feladat optimális megoldásának folytonos (\mathbf{x}) részét a dualitás tétel alapján (PRÉKOPA [55]) az utolsóként megoldott (6.8) feladat

optimális megoldásához tartozó szimplex tábla jellemzőiből előállíthatjuk, nincs szükség újabb lineáris programozási feladat megoldására (HOFFER [31]).

További egyszerűsítés az, hogy a (6.8) feladatok megoldására az első iterációban kétfázisú szimplex módszert alkalmazunk (PRÉKOPA [55]); a többi iterációban a feladat megoldását a második fázisban kezdjük, induló megengedett megoldásként a megelőző iterációban nyert optimális megoldás szolgál.

	változók száma:		
	(≤ 30)	($= 2$)	(≤ 35)
feltételek száma ≤ 35		1 -1	E
		1 -1	
		1 -1	
		...	
		...	
		1 -1	

16. ábra

A feszültségellenőrző periódusok feladatának megoldása során fellépő (6.8) lineáris programozási feladat két független blokkból áll. Ezért megoldását két független feladatra bontjuk (a felbontás után kapott két lineáris programozási feladat mátrixát lásd a 16. és 17. ábrán).

	változók száma:			
	(≤ 30)	(≤ 30)	(≤ 30)	(≤ 30)
felt. száma ≤ 30	A ₁ ^T	- A ₁ ^T	A ₂ ^T	E

17. ábra

A feszültséges periódusok feladatában megoldandó LP két független blokkja

A normál periódusok feladatának megoldása során fellépő (6.8) lineáris programozási feladat megoldására vonatkozó egy további ötlet felhasználása folyamatban van; a HOFFER [32]-ben javasolt algoritmus alkalmazása jelentősen csökkentené a napi menetrend előállításához szükséges számítási időt.

Megjegyezzük még, hogy az **F** mátrix szerepe a dekompozíciós eljárás során különleges. Kizárólag mátrix-szorzási műveletekre használjuk azt; jobbról szorzunk vele a (6.6) relaxált feladatok feltételeinek előállításakor, balról pedig a (6.8) feladatok célfüggvény együtthatóinak kiszámításához. Minthogy mind a normál, mind a feszültségellenőrző periódusok feladatában az üzemmód változókhoz tartozó mátrixok jól struktúráltak — kivéve az ácterhelési feltételek részmatrixát — a velük való szorzási szabály könnyen leírható. Ezt a tényt kihasználva a programban nem töltöttünk fel egy mátrixot az elemekkel, hanem a szorzási műveleteket a speciális szerkezetnek megfelelően végezzük.

Az üzemmód változókra vonatkozó alábbi feltételeket (amelyek az Y halmaz tulajdonságait írják le) a (6.6) relaxált feladatainak megoldására alkalmazott leszám-lálási algoritmuson belül a következőképpen érvényesítjük:

— egy megelőző és egy rákövetkező periódus üzemmód rendszernek a figyelem-bevételét a megfelelő üzemmód változók 0 vagy 1 szinten való rögzítésével;

— az adott periódusra vonatkozó további előírások betartását: bizonyos üzem-módok kötelező üzemeltetését vagy kötelező kikapcsolását hasonlóan az üzemmód-változók rögzítésével;

— az üzemmód változók speciális logikáját leíró (5.19) feltételek teljesítése a leszám-lálásbeli rögzítések és lekötések összes következményének megvizsgálásával és a megfelelő változók rögzítésével történik. Pl. ha valamely üzemmód változó 1 értéket kap a leszám-lálás során, akkor az ugyanahhoz az erőműhöz tartozó kisebb indexű üzemmód változókat 1 értéken lerögzítjük. Ha ugyanez a változó 0 értéket kapna, akkor a megfelelő nagyobb indexű üzemmódváltozókat rögzítjük le 0 szinten.

Függelék

Villamosenergia-rendszerek átviteli hálózatáról

F1. A villamos hálózatok matematikai modellje

Ebben a részben egy villamos hálózatokra vonatkozó matematikai modellt építünk fel. Ehhez felhasználunk bizonyos szakirodalmat, így elsősorban a [39], [60], [69] műveket, azoktól azonban eltérünk, mert az átviteli hálózat matematikai modell-jének konstrukciója megkívánja az ismert tényeknek a szokásostól eltérő rendszer szerinti szisztematikus tárgyalását, továbbá bizonyos fogalmak új megvilágításba helyezését. Így például az általunk bevezetett admittancia-transzformáció a szakiroda-lomban szereplő transzformációhoz képest lényegesen egyszerűbb, a fizikai szemlélet-hez nézetünk szerint jobban igazodó tárgyalásmódot tesz lehetővé.

A villamos hálózatot irányított gráffal modellezzük, ezért a teljesség kedvéért röviden összefoglaljuk a szükséges gráfelméleti alapismereteket.

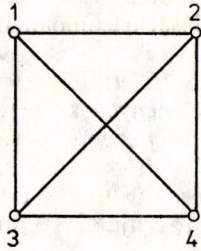
Tekintsünk egy \mathcal{N} véges halmazt, továbbá ennek bizonyos elempárjaiból alkot-sunk meg egy további \mathcal{A} halmazt. Az \mathcal{N} és \mathcal{A} halmazokat együtt gráfnak, az \mathcal{N} halmaz elemeit csúcsoknak, az \mathcal{A} halmaz elemeit éleknek nevezzük. Az \mathcal{A} halmazban foglalt elempárok lehetnek rendezettek (amikor megmondjuk, hogy a két elem közül melyik az első és melyik a második), vagy nem rendezettek. Az előbbi esetben az éleket irányítottaknak, az utóbbi esetben nem irányítottaknak (irányítatlanoknak) nevezzük. A nem irányított élekkel bíró gráfot irányítatlan, az irányított élekkel bíró pedig irányított gráfnak nevezzük.

Ez utóbbinál a könnyebb megkülönböztethetőség kedvéért az élek halmazát \mathcal{A}^D -vel jelöljük. A fenti módon megadott gráfok jelölésére az irányítatlan esetben az $(\mathcal{N}, \mathcal{A})$, az irányított esetben pedig az $(\mathcal{N}, \mathcal{A}^D)$ szimbólumokat alkalmazzuk.

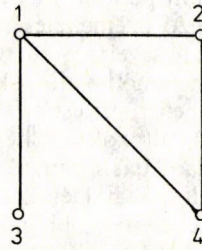
Először irányítatlan gráfokkal foglalkozunk. Ha az \mathcal{N} halmaznak n eleme van, akkor a gráf csúcsainak jelölésére az $1, \dots, n$ számokat is alkalmazhatjuk. Ekkor minden él egy $\{i, j\}$ számpár, pontosabban az $\mathcal{N} = \{1, \dots, n\}$ halmaz egy két elemű részhalmaza. Azt mondjuk, hogy az $\{i, j\}$ él az i és j csúcsokat köti össze.

Világos, hogy ha a csúcsok száma n , akkor az élek száma legfeljebb $n(n-1)/2$ lehet.

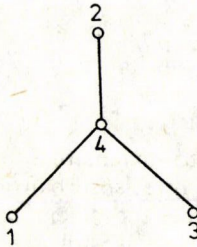
Egy gráfot oly módon szemléltetünk, hogy a síkban felvesszünk n számú pontot, ezek jelentik a gráf csúcsait és vonallal összekötjük az i, j pontokat, ha $\{i, j\} \in \mathcal{A}$. Az alábbi ábrákon négy gráfot láthatunk:



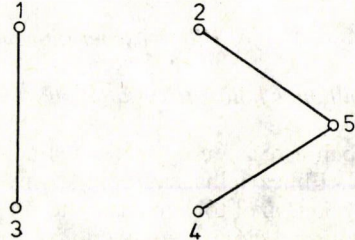
F1. ábra



F2. ábra



F3. ábra



F4. ábra

Az alábbiakban néhány fontos fogalmat ismertetünk, majd néhány tételt bizonyítunk be.

Részgráf. A $\mathcal{G}' = (\mathcal{N}', \mathcal{A}')$ gráfot a $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ gráf részgráfjának nevezzük, ha $\mathcal{N}' \subseteq \mathcal{N}$ és $\mathcal{A}' \subseteq \mathcal{A}$.

Út. Ha s, i_1, \dots, i_k, t az \mathcal{N} halmaz különböző elemei, akkor az élek $\{s, i_1\}, \{i_1, i_2\}, \dots, \{i_{k-1}, i_k\}, \{i_k, t\}$ sorozatát az s, t csúcsokat összekötő útnak nevezzük.

Hurok. (Körút) Ha s, i_1, \dots, i_k az \mathcal{N} halmaz különböző elemei, akkor az élek $\{s, i_1\}, \{i_1, i_2\}, \dots, \{i_{k-1}, i_k\}, \{i_k, s\}$ sorozatát huroknak nevezzük.

Összefüggő gráf. Ha a gráf bármely két különböző csúcsa összeköthető úttal, akkor azt mondjuk, hogy a gráf összefüggő. (Az F4. ábrán látható gráf nem összefüggő, az F1., F2., F3. ábrán ábrázolt gráfok összefüggők.)

Fa. Egy gráfot fának nevezzük, ha összefüggő és nincs benne hurok. (Az F3. ábrán látható gráf fa.)

Feszítő fa. Az $(\mathcal{N}, \mathcal{A})$ gráf $(\mathcal{N}', \mathcal{A}')$ részgráfját feszítő fának nevezzük, ha $(\mathcal{N}', \mathcal{A}')$ fa és $\mathcal{N} = \mathcal{N}'$.

Izolált pont. Ha az i csúcs egyetlen élre sem illeszkedik, akkor azt izolált csúcsnak nevezzük.

Végpont (terminális pont). Ha az i csúcs pontosan egy élre illeszkedik, akkor azt a gráf végpontjának, másnéven terminális pontjának nevezzük.

F1.1. TÉTEL. Egy n csúccsal bíró fának $n-1$ éle van.

Bizonyítás. Az állítás az $n=1$ esetben nyilvánvalóan igaz (egyetlen csúcsból álló gráf összefüggő és hurokmentes, tehát fa).

Az $n \geq 2$ esetben először bebizonyítjuk, hogy a fának van terminális pontja. Ha ugyanis tekintjük a gráfban konstruálható leghosszabb utat, akkor az a két pont, melyeket ez az út összeköt, terminális pont, hiszen ellenkező esetben az utat meg lehetne hosszabbítani.

Ezek után a tételt az $n \geq 2$ esetben teljes indukcióval bizonyítjuk. Az $n=2$ esetben a tétel állítása nyilvánvalóan igaz. Feltéve, hogy az állítás igaz minden n csúccsal bíró fára, tekintsünk egy $n+1$ csúccsal bíró fát, ahol $n \geq 2$. Ennek egy tetszőleges terminális pontját és a hozzá csatlakozó élt elhagyva, egy n csúccsal bíró fát kapunk, melynek az indukciós feltevés szerint $n-1$ éle van. A most elhagyott éllel együtt tehát összesen n él van, és így az állítást $n+1$ csúcs esetére bebizonyítottuk. Ezzel az F1.1. tétel bizonyítása készen van.

F1.2. TÉTEL. Egy fa minden csúcspárja pontosan egy úttal köthető össze.

Bizonyítás. Egy út létezését a fa definíciója is megkívánja. Ha volna olyan csúcspár, melyet két különböző úttal össze lehet kötni, akkor amint azt könnyű belátni, volna hurok is a gráfban. Ám ez nem lehetséges, mert a fa hurokmentes. Ezzel a bizonyítás készen van.

F1.3. TÉTEL. Ha egy gráf minden csúcspárja pontosan egy úttal köthető össze, akkor a gráf fa.

Bizonyítás. Minthogy a gráf a feltétel szerint összefüggő, csak a hurokmentességet kell belátni. Ez azonban egyszerűen következik, hiszen a hurok két tetszőleges csúcsa a hurokban levő élek felhasználásával két különböző módon úttal összeköthető. Ezzel az F1.3. tételt bebizonyítottuk.

A fenti tételeket célszerű az alábbi, egyetlen tételben összefoglalni.

F1.4. TÉTEL. Egy n csúccsal bíró \mathcal{G} gráf esetén az alábbi állítások ekvivalensek:

- a \mathcal{G} gráf fa;
- a \mathcal{G} gráf minden csúcspárja egyértelműen összeköthető úttal;
- \mathcal{G} összefüggő és $n-1$ éle van;
- \mathcal{G} hurokmentes és $n-1$ éle van.

F1.5. TÉTEL. Ha a $\mathcal{G}=(\mathcal{N}, \mathcal{A})$ gráf fa és $\{i, j\} \notin \mathcal{A}$, akkor a $\mathcal{G}_1=(\mathcal{N}, \mathcal{A} \cup \{i, j\})$ gráfban pontosan egy hurok van.

Bizonyítás. A \mathcal{G}_1 gráfban biztosan van hurok, ugyanis a csúcsainak a száma ugyanannyi, mint a \mathcal{G} gráf esetén, ám eggyel több éle van; a \mathcal{G}_1 gráf tehát nem fa, viszont összefüggő, amiből következik, hogy van benne hurok. Azt kell csak belát-

nunk, hogy nincs két hurok \mathcal{G}_1 -ben. Indirekt bizonyítással élve, tegyük fel, hogy \mathcal{G}_1 -ben van két hurok. Ekkor az $\{i, j\}$ él mindkettőben benne kell, hogy legyen, hiszen \mathcal{G} -ben nincsen hurok. Mindkét hurokból az $\{i, j\}$ élt elhagyva az i és j csúcsokat összekötő két különböző utat kapunk, ami az F1.2. tétel szerint nem lehetséges.

F1.6. TÉTEL. Ha a $\mathcal{G}=(\mathcal{N}, \mathcal{A})$ gráf összefüggő, akkor tartalmaz feszítő fát.

Bizonyítás. Konstruktív bizonyítást alkalmazunk. Legyen $\mathcal{N}_1=\{i_1\}$, $\mathcal{A}_1=\emptyset$, ahol $i_1 \in \mathcal{N}$ tetszőleges, ekkor $\mathcal{G}_1=(\mathcal{N}_1, \mathcal{A}_1)$ részgráfja \mathcal{G} -nek és triviálisan fa. Tételezzük fel, hogy a \mathcal{G} gráf $\mathcal{G}_k=(\mathcal{N}_k, \mathcal{A}_k)$ részgráfja adott és ez fa, $1 \leq k \leq n-1$. Ekkor \mathcal{G}_{k+1} -et a következő módon konstruáljuk: Mivel \mathcal{G} összefüggő, ezért van olyan $i_p \in \mathcal{N}_k$ és $i_q \in \mathcal{N} \setminus \mathcal{N}_k$, melyre $\{i_p, i_q\} \in \mathcal{A}$. Legyen $\mathcal{N}_{k+1}=\mathcal{N}_k \cup \{i_q\}$ és $\mathcal{A}_{k+1}=\mathcal{A}_k \cup \{i_p, i_q\}$. Ekkor világos, hogy $\mathcal{G}_{k+1}=(\mathcal{N}_{k+1}, \mathcal{A}_{k+1})$ újra fa lesz és így $\mathcal{G}_n=(\mathcal{N}_n, \mathcal{A}_n)$ feszítő fa. Ezzel a bizonyítást befejeztük.

Az irányítatlan gráf éleit az eddigiek során az illető él végpontjaiból alkotott csúcspárral azonosítottuk. A villamos hálózatok modellezéséhez azonban olyan irányítatlan gráf fogalomra van szükség, amelyben két csúcspontot egynél több él is összeköthet, így az élek a végpontjaiból alkotott csúcspárral már nem lesznek azonosíthatók. Ezért az irányítatlan gráf definícióját általánosabban fogalmazzuk meg, a továbbiak során irányítatlan gráfon a következőket fogjuk érteni:

Tekintsük az \mathcal{N} , \mathcal{A} véges halmazokat és rendeljünk hozzá \mathcal{A} minden egyes eleméhez az \mathcal{N} halmaz elemeiből alkotott egy-egy rendezetlen elempárt, pontosabban \mathcal{N} valamely kételemű részhalmazát. Az \mathcal{N} , \mathcal{A} halmazokat a fenti hozzárendeléssel együtt irányítatlan gráfnak nevezzük és az egyszerűség kedvéért továbbra is az $(\mathcal{N}, \mathcal{A})$ szimbólummal jelöljük. Az \mathcal{N} halmaz elemeit csúcsoknak, az \mathcal{A} halmaz elemeit éleknek nevezzük.

Legyen \mathcal{N} elemszáma n , \mathcal{A} elemszáma m . A gráf csúcsainak jelölésére itt is az $1, \dots, n$ számokat alkalmazhatjuk, a gráf éleit jelölje rendre e_1, e_2, \dots, e_m . Ha $e_k \in \mathcal{A}$ és az ehhez az élhez hozzárendelt elempár $\{i, j\}$, ezt az $e_k \rightarrow \{i, j\}$ szimbólummal jelöljük. Amennyiben $e_k \rightarrow \{i, j\}$ és $e_l \rightarrow \{i, j\}$, $k \neq l$ teljesül, akkor az e_k, e_l éleket párhuzamos éleknek nevezzük.

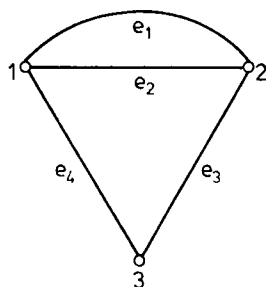
Világos, hogy az indulásként definiált irányítatlan gráf mostani definíciónkból speciális esetként adódik, amennyiben a definícióban szereplő hozzárendelésről kikötjük, hogy különböző élekhez különböző csúcspárokat rendeljen.

A gráf szemléltetése az eddigiekhez hasonlóan történhet. A síkon felvett n számú pont reprezentálja a gráf csúcsait, és az $\{i, j\}$ pontokat annyi vonallal kötjük össze, ahány élhez az $\{i, j\}$ számpár hozzá van rendelve. Az élek mellé felírjuk az él azonosítóját. Az F5. ábrán látható gráfban e_1, e_2 párhuzamos élek.

A részgráf, út, hurok definíciója a következőképpen módosul:

Részgráf. A $\mathcal{G}'=(\mathcal{N}', \mathcal{A}')$ gráfot a $\mathcal{G}=(\mathcal{N}, \mathcal{A})$ gráf részgráfnak nevezzük, ha $\mathcal{N}' \subseteq \mathcal{N}$, $\mathcal{A}' \subseteq \mathcal{A}$, továbbá az \mathcal{A}' halmaz elemeihez ugyanazokat az elempárokat rendeljük a \mathcal{G}' gráfban, mint a \mathcal{G} gráfban.

Út. Ha s, i_1, \dots, i_k, t az \mathcal{N} halmaz különböző elemei, akkor az élek $e_{j_1}, \dots, e_{j_{k+1}}$ sorozatát az s, t csúcsokat összekötő útnak nevezzük, amennyiben $e_{j_1} \rightarrow \{s, i_1\}$, $e_{j_2} \rightarrow \{i_1, i_2\}$, \dots , $e_{j_k} \rightarrow \{i_{k-1}, i_k\}$, $e_{j_{k+1}} \rightarrow \{i_k, t\}$ érvényes.



F5. ábra

Hurok. Ha s, i_1, \dots, i_k az \mathcal{N} halmaz különböző elemei, akkor az élek $e_{j_1}, \dots, e_{j_{k+1}}$ sorozatát huroknak nevezzük, amennyiben $e_{j_1} \rightarrow \{s, i_1\}$, $e_{j_2} \rightarrow \{i_1, i_2\}$, \dots , $e_{j_k} \rightarrow \{i_{k-1}, i_k\}$, $e_{j_{k+1}} \rightarrow \{i_k, s\}$ teljesül.

Az indulásként definiált irányítatlan gráfra vonatkozó összes többi definíciót — a fentiekben módosított részgráf, út, hurok fogalmakat véve alapul — szó szerint megismételve adódnak a megfelelő fogalmak az általánosabban definiált irányítatlan gráfra vonatkozóan. Az Olvasó könnyen ellenőrizheti, hogy valamennyi eddigi tétel érvényes lesz az általánosabban tekintett gráfok esetében is, ugyanazok a bizonyítások alkalmazhatók, helyenkénti értelemszerű átfogalmazással.

Ezután irányított gráfokkal foglalkozunk. Ha az $(\mathcal{N}, \mathcal{A}^D)$ irányított gráfnak n számú csúcsa van, ezek jelölésére most is alkalmazhatjuk az $1, \dots, n$ számokat. Ekkor minden élnek egy (i, j) rendezett számpár felel meg. Feltételezzük, hogy $\mathcal{A}^D \subseteq \mathcal{N} \times \mathcal{N}$, ami azt jelenti, hogy mindegyik rendezett számpár legfeljebb egy élhez van hozzárendelve. Következésképpen az élek jelölésére a megfelelő rendezett számpárokat alkalmazhatjuk. Legyen az élek száma m , és lássuk el az éleket is sorszámmal. Feltételezzük, hogy $(\mathcal{N}, \mathcal{A}^D)$ nem tartalmaz hurokért, azaz $(i, i) \notin \mathcal{A}^D$, $i = 1, \dots, n$ teljesül.

Jelölje $(\mathcal{N}, \mathcal{A})$ az $(\mathcal{N}, \mathcal{A}^D)$ gráfnak megfelelő irányítatlan gráfot, amely az $(\mathcal{N}, \mathcal{A}^D)$ gráfból úgy kapható, hogy annak éleit irányítatlanoknak tekintjük.

Az egyszerűbb jelölésmód kedvéért a továbbiakban megengedjük a következőt: Ha $\mathbf{d} \in \mathbb{R}^m$ (vagy $\mathbf{d} \in \mathbb{C}^m$) komponensei él-jellemzők, és az l -edik élt az (i, j) számpár reprezentálja, akkor d_l mellett megengedjük a $d_{i,j}$ jelölést is a \mathbf{d} vektor l -edik komponensére. Tekintsük továbbá $(\mathcal{N}, \mathcal{A})$ azon éleit, amelyek valamelyik végpontja az i -edik csomópont, $i = 1, \dots, n$, és jelöljük $J(i)$ -vel a másik végpontok sorszámainak halmazát. Az ezen éleknek az $(\mathcal{N}, \mathcal{A}^D)$ gráfban megfelelő élek halmazát az i -edik csomópont-hoz illeszkedő élek halmazának fogjuk nevezni. Jelölje továbbá $i(l)$ az i -edik csomópont-hoz illeszkedő l sorszámu él másik végpontjának sorszámat.

DEFINÍCIÓ. Az $(\mathcal{M}, \mathcal{B}^D)$ irányított gráfot az $(\mathcal{N}, \mathcal{A}^D)$ irányított gráf részgráfiának nevezzük, ha $\mathcal{M} \subseteq \mathcal{N}$ és $\mathcal{B}^D \subseteq \mathcal{A}^D$ teljesül.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ irányított gráf összefüggő, ha $(\mathcal{N}, \mathcal{A})$ összefüggő.

A továbbiakban kizárólag összefüggő gráfokkal foglalkozunk, azaz feltételezzük, hogy $(\mathcal{N}, \mathcal{A}^D)$ összefüggő, irányított gráf.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ gráf $(\mathcal{N}, \mathcal{F}^D)$ részgráfja feszítő fa, ha az $(\mathcal{N}, \mathcal{A})$ irányítatlan gráfban $(\mathcal{N}, \mathcal{F})$ feszítő fa.

Az $\mathcal{A}^D \setminus \mathcal{F}^D$ elemeit kötőéleknek nevezzük, a továbbiakban jelölje ezek számát k . Az F1.1. tétel következményeként az $(\mathcal{N}, \mathcal{F}^D)$ éleinek száma $n-1$ és $k=m-n+1$ teljesül.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ gráf $(\mathcal{N}_H, \mathcal{H}^D)$ részgráfja hurok, ha az $(\mathcal{N}, \mathcal{A})$ gráfban $(\mathcal{N}_H, \mathcal{H})$ hurok.

A hurkokhoz irányítást is rendelhetünk, ami intuitíve egy „körüljárás” rögzítését jelenti $(\mathcal{N}_H, \mathcal{H})$ -ra. Ha a hurokban kettőnél több csúcs szerepel, akkor ez azt jelenti, hogy ezen csúcsokra vonatkozóan egy ciklikus rendezést rögzítünk. Ekkor ha $(i, l) \in \mathcal{H}^D$, az (i, l) élt a hurokra nézve pozitívan irányítottan nevezzük, ha a ciklikus rendezésben a csúcsok sorrendje i, l ; ellenkező esetben (i, l) a hurokra nézve negatívan irányított. Ha a hurokban csak két csúcs szerepel, akkor $\mathcal{H}^D = \{(i, l), (l, i)\}$ alakú és a „körüljárás” pontos definíciója a fentiekhez hasonlóan egyszerűen megfogalmazható.

DEFINÍCIÓ. Legyen $\mathcal{N}_1, \mathcal{N}_2$ az \mathcal{N} halmaz egy particiója, azaz teljesüljön $\mathcal{N}_1 \cup \mathcal{N}_2 = \mathcal{N}$ és $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$. Az ehhez a particióhoz tartozó vágásnak nevezzük, és $C(\mathcal{N}_1, \mathcal{N}_2)$ -vel jelöljük ebben a részben, \mathcal{A}^D következő részhalmazát:

$$C(\mathcal{N}_1, \mathcal{N}_2) = \{(i, l) | (i, l) \in \mathcal{A}^D, i \in \mathcal{N}_1, l \in \mathcal{N}_2, \text{ vagy } i \in \mathcal{N}_2, l \in \mathcal{N}_1\}.$$

A vágásokhoz irányítást is rendelhetünk. Ez azt jelenti, hogy $\mathcal{N}_1, \mathcal{N}_2$ egyik lehetséges sorrendjét rögzítjük, azaz pl. az $(\mathcal{N}_1, \mathcal{N}_2)$ halmazpárt rendezett párként kezeljük. Ha $(i, l) \in C(\mathcal{N}_1, \mathcal{N}_2)$, akkor ezt az élt vágásra nézve pozitívan irányítottan tekintjük, ha $i \in \mathcal{N}_1$, és negatívan irányítottan, ha $i \in \mathcal{N}_2$ teljesül.

DEFINÍCIÓ. Az l -edik csúcsához tartozó csúcs-vágásnak nevezzük $l=1, \dots, n$ -re a következő vágást: $C(\{l\}, \mathcal{N} \setminus \{l\})$. Látható, hogy $C(\{l\}, \mathcal{N} \setminus \{l\})$ az l -edik csúcs-hoz illeszkedő élek halmaza.

Az egyszerűbb tárgyalásmód kedvéért kiválasztunk az $(\mathcal{N}, \mathcal{A}^D)$ gráfból egy $(\mathcal{N}, \mathcal{F}^D)$ feszítő fát, melyet a továbbiakban rögzítettnek tekintünk.

DEFINÍCIÓ. Az $\mathcal{A}^D \setminus \mathcal{F}^D$ kötőélekhez tartozó kötő-hurkok a következőképpen adódó hurkok lesznek: Legyen $(i, l) \in \mathcal{A}^D \setminus \mathcal{F}^D$. Az F1.5. tétel következményeként az $(\mathcal{N}, \mathcal{F}^D \cup \{(i, l)\})$ gráf egy és csak egy hurkot tartalmaz, melyet az (i, l) él által generált kötő-huroknak nevezünk.

Az $(\mathcal{N}, \mathcal{F}^D)$ fához tartozó kötő-hurkok száma a fentiek szerint k .

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{F}^D)$ fa éleihez tartozó fa-vágásokon a következőképp adódó vágásokat értjük. Legyen $(i, l) \in \mathcal{F}^D$. A fából törölve az (i, l) élt, az két diszjunkt fára esik szét, ilymódon a két fa csúcspontjait tekintve \mathcal{N} egy particiója adódik. Az ehhez a particióhoz tartozó $(\mathcal{N}, \mathcal{A}^D)$ -beli vágást nevezzük az (i, l) él által generált fa-vágásnak.

A definícióból és az F1.1. tételből látható, hogy az $(\mathcal{N}, \mathcal{F}^D)$ fához tartozó fa-vágások száma $n-1$.

A továbbiakban szükségünk lesz az $(\mathcal{N}, \mathcal{A}^D)$ gráf struktúráját és a vágások, ill. hurkok irányítását leíró különféle incidencia mátrixokra. Feltételezzük, hogy a gráf-

hoz tartozó vágások és hurkok sorszámozva vannak, és az így adódó rendezésnek megfelelően a fa-vágások, kötő-hurkok, valamint a csúcs-vágások külön-külön is sorszámokkal vannak ellátva. Ezek után az incidencia mátrixok definíciói rendre a következők:

Q: A gráfhoz tartozó vágásokat és azok valamely rögzített irányítását leíró vágás, él incidencia mátrix. Mérete $p \times m$, ahol p az $(\mathcal{N}, \mathcal{A}^D)$ gráfhoz tartozó vágások száma. \hat{Q} definíciója a következő:

$$(F1.1) \quad \hat{Q}_{il} = \begin{cases} 1, & \text{ha az } l\text{-edik él szerepel az } i\text{-edik vágásban és a vágásra nézve} \\ & \text{pozitív irányítású;} \\ -1, & \text{ha az } l\text{-edik él szerepel az } i\text{-edik vágásban és a vágásra nézve} \\ & \text{negatív irányítású;} \\ 0, & \text{különben.} \end{cases}$$

Q: Redukált vágás, él incidencia mátrix, mérete $(n-1) \times m$. Definíciójához feltételezzük, hogy a vágásokat a következő irányítással láttuk el: a fa-vágások irányítása olyan, hogy arra nézve a vágást generáló él pozitív irányítású, a többi vágás irányítása tetszőlegesen rögzített. Ekkor \mathbf{Q} a $\hat{\mathbf{Q}}$ mátrixból az $n-1$ számú fa-vágásnak megfelelő sor kiválasztásával adódó részmátrix.

Â: Csúcs, él incidencia mátrix, $n \times m$ méretű. Feltételezzük, hogy $(\mathcal{N}, \mathcal{A}^D)$ vágásait a következő irányítással láttuk el: a csúcs-vágások irányítása olyan, hogy arra nézve a csúcsból kifelé mutató élek pozitív irányításúak, a többi vágás irányítása tetszőlegesen rögzített. Ekkor $\hat{\mathbf{A}}$ a $\hat{\mathbf{Q}}$ mátrix n számú csúcs-vágásához tartozó sorának megfelelő részmátrixa. Mivel a csúcs, él incidencia-mátrixra a későbbiekben gyakran lesz szükségünk, explicit alakját is megadjuk:

$$(F1.2) \quad \hat{A}_{il} = \begin{cases} 1, & \text{ha az } l\text{-edik él illeszkedik az } i\text{-edik csúcsra és a csúcs az él kez-} \\ & \text{dőpontja;} \\ -1, & \text{ha az } l\text{-edik él illeszkedik az } i\text{-edik csúcsra és a csúcs az él vég-} \\ & \text{pontja;} \\ 0, & \text{különben.} \end{cases}$$

A: Redukált csúcs, él incidencia mátrix, $(n-1) \times m$ méretű. A mátrix $\hat{\mathbf{A}}$ -ból annak valamely sora törlésével keletkezik, így $\hat{\mathbf{A}}$ -hoz n számú redukált csúcs, él incidencia mátrix rendelhető.

Ĥ: Az $(\mathcal{N}, \mathcal{A}^D)$ gráfhoz tartozó hurkokat és azok valamely rögzített irányítását leíró hurok, él incidencia mátrix. Mérete $h \times m$, ahol h a hurkok száma a gráfban. Definíciója az alábbi:

$$(F1.3) \quad \hat{B}_{il} = \begin{cases} 1, & \text{ha az } i\text{-edik hurokban szerepel az } l\text{-edik él és a hurokra nézve} \\ & \text{pozitív irányítású;} \\ -1, & \text{ha az } i\text{-edik hurokban szerepel az } l\text{-edik él és a hurokra nézve} \\ & \text{negatív irányítású;} \\ 0, & \text{különben.} \end{cases}$$

B: Redukált hurok, él incidencia mátrix, $k \times m$ méretű. Definíciójához a hurkokat a következőképpen irányítjuk: A kötő-hurkok irányítása olyan, hogy arra nézve a hurkot generáló él pozitív irányítású, a többi hurok irányítása tetszőlegesen rögzített. Ekkor \mathbf{B} a $\hat{\mathbf{B}}$ mátrix (kötő-hurkoknak) megfelelő soraiból adódó rész-mátrix.

Az egyszerűbb jelölésmód kedvéért tételezzük fel, hogy a gráf éleinek sorszámozása a rögzített $(\mathcal{N}, \mathcal{F}^D)$ fára vonatkozó kötőélekkel kezdődik. Ekkor \mathbf{B} és \mathbf{Q} a következő ún. normál alakot ölti:

$$(F1.4) \quad \mathbf{B} = (\mathbf{E}_k, \mathbf{F}); \quad \mathbf{Q} = (\mathbf{Q}', \mathbf{E}_{n-1}),$$

ahol \mathbf{E}_k , \mathbf{E}_{n-1} a megfelelő méretű egység mátrixok.

A következőkben szükségünk lesz az alábbi tételre, melynek bizonyítása megtalálható pl. [60]-ban.

F1.7. TÉTEL. Az $(\mathcal{M}, \mathcal{B}^D)$ irányított gráf akkor és csak akkor tartalmaz hurkot, ha az $\hat{\mathbf{A}}$ csúcs, él incidencia mátrixának oszlopai lineárisan összefüggők.

A különböző incidencia mátrixok rangjáról szól a következő tétel. Megfogalmazásához vezessük be a mátrixok rangjára az $r(\mathbf{D})$ jelölést, ahol \mathbf{D} valamely mátrix.

F1.8. TÉTEL. Az $(\mathcal{N}, \mathcal{A}^D)$ összefüggő gráf incidencia mátrixaira érvényesek az alábbi állítások:

- (i) $r(\mathbf{Q}) = n - 1$, $r(\mathbf{B}) = k$.
- (ii) $r(\hat{\mathbf{A}}) = n - 1$.
- (iii) $\hat{\mathbf{A}}$ bármelyik sorát törölve, az adódó \mathbf{A} redukált csúcs, él incidencia mátrix sorai lineárisan függetlenek.

Bizonyítás. Az (i) állítás (F1.4) alapján triviális. Az (ii) bizonyításához először is megjegyezzük, hogy $\hat{\mathbf{A}}$ sorvektorait összeadva a zérus vektor adódik. Így $\hat{\mathbf{A}}$ sorai lineárisan összefüggők, azaz $r(\hat{\mathbf{A}}) \leq n - 1$. Ugyanakkor $\hat{\mathbf{A}}$ oszlopvektorai közül kiválasztható $n - 1$ lineárisan független. Ehhez elegendő egy — az F1.6. tétel szerint létező — feszítő fának megfelelő oszlopvektorokat kiválasztani, melyek az F1.7. tétel szerint lineárisan függetlenek. Végül (iii) bizonyítása abból könnyen adódik, — hogy az (ii) rész bizonyításának kezdetén tett megjegyzésünk miatt — $\hat{\mathbf{A}}$ bármely sora kifejezhető a többi lineáris kombinációjaként. Ezzel a tétel bizonyítását befejeztük.

A következő tétel egy ortogonalitási relációt fogalmaz meg.

F1.9. TÉTEL. Az $(\mathcal{N}, \mathcal{A}^D)$ gráf vágásainak és hurkainak tetszőleges irányítása mellett érvényes az alábbi reláció:

$$(F1.5) \quad \hat{\mathbf{Q}}\hat{\mathbf{B}}^T = \mathbf{0}.$$

Bizonyítás. Válasszuk ki a $\hat{\mathbf{Q}}$ mátrix s -edik és a $\hat{\mathbf{B}}$ mátrix t -edik sorát. Ez azt jelenti, hogy az s -edik vágást és a t -edik hurkot tekintjük. Ha a vágásnak és a huroknak nincsenek közös élei, akkor a skalár szorzat triviálisan zérus. Egyébként egyszerű belátni, hogy a vágásnak és a huroknak páros sok közös éle van. A hurok irányításá-

nak megfelelő körüljárást véve a közös éleknek a hurokra és a vágásra vonatkozó irányítása felváltva egyező és ellenkező, amiből az állítás már könnyen következik.

A továbbiak során jelölje \mathbf{C}^m a komplex szám \mathbf{C} feletti lineáris terét, a szósorozatos skalár szorzattal: $\mathbf{a}, \mathbf{b} \in \mathbf{C}^m$ esetén $\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=1}^m a_i b_i^*$, ahol a $*$ komplex konjugáltat jelent.

Ha \mathbf{D} altere \mathbf{C}^m -nek, jelölje a dimenzióját $\dim \mathbf{D}$.

Mivel az F1.8., F1.9. tételekben szereplő mátrixok elemei valós számok, ezért a rangokra vonatkozó állítások \mathbf{C}^m -et alapul véve is érvényesek, és az (F1.5) ortogonalitási reláció \mathbf{C}^m -beli skalár-szorzatokkal is megfogalmazható.

Bevezetjük a következő jelöléseket:

Jelöljük \mathbf{C}_Q -val \mathbf{C}^m -nek a $\hat{\mathbf{Q}}$ sorvektorai által felfeszített alterét, \mathbf{C}_B pedig jelölje a $\hat{\mathbf{B}}$ sorvektorai által felfeszített alteret. Mivel a vágások, ill. hurkok különböző irányításaihoz tartozó $\hat{\mathbf{Q}}$, ill. $\hat{\mathbf{B}}$ mátrixok sorai legfeljebb (-1) -es szorzóban különböznek, ezért a \mathbf{C}_Q , ill. a \mathbf{C}_B nem függ a vágások, illetve a hurkok irányításától.

F1.10. TÉTEL. Érvényesek az alábbi állítások:

- (i) $\dim \mathbf{C}_Q = n - 1$;
- (ii) $\dim \mathbf{C}_B = k$;
- (iii) \mathbf{C}_Q és \mathbf{C}_B egymás ortogonális kiegészítő alterei.

Bizonyítás. Mivel a vágások és hurkok alkalmas irányítása mellett $\hat{\mathbf{Q}}$ sorai között megtalálhatók a \mathbf{Q} mátrix, $\hat{\mathbf{B}}$ sorai között a \mathbf{B} mátrix sorai, ezért az F1.8. tétel felhasználásával az alábbi adódik:

$$(F1.6) \quad \dim \mathbf{C}_Q \cong n - 1, \quad \dim \mathbf{C}_B \cong k.$$

Másrészt (F1.5) felhasználásával könnyen belátható, hogy \mathbf{C}_Q és \mathbf{C}_B ortogonális alterei \mathbf{C}^m -nek. Így

$$(F1.7) \quad \dim \mathbf{D}_Q + \dim \mathbf{C}_B \leq m$$

adódik. Ebből az $m = n - 1 + k$ összefüggés felhasználásával megkaptuk (i) és (ii) bizonyítását. Az alterek ortogonalitása és (i) valamint (ii) implikálják (iii)-t. Ezzel a bizonyítást befejeztük.

Tételünk felhasználásával teljessé tehetjük az incidencia mátrixok rangjára vonatkozó állítások sorozatát.

F1.11. TÉTEL. Az $(\mathcal{N}, \mathcal{A}^D)$ összefüggő gráfra vonatkozóan a vágások, illetve a hurkok tetszőleges irányítása mellett $r(\hat{\mathbf{Q}}) = n - 1$ és $r(\hat{\mathbf{B}}) = k$ érvényes.

Bizonyítás. Az F1.10. tétel alapján triviális.

Megjegyezzük, hogy az (F1.5) ortogonalitási relációt felhasználva az (F1.4) összefüggésekben $\mathbf{Q}' = -\mathbf{F}^T$ érvényes, így \mathbf{B} és \mathbf{Q} a következő alakban írható:

$$(F1.8) \quad \mathbf{B} = (\mathbf{E}_k, \mathbf{F}), \quad \mathbf{Q} = (-\mathbf{F}^T, \mathbf{E}_{n-1}).$$

Az éleknek az (F1.4) összefüggéseknél rögzített sorszámozásából adódik a C^m -beli vektorok komponenseinek következő partíciója:

$$\text{Ha } \mathbf{d} \in C^m, \text{ akkor } \mathbf{d} = \begin{pmatrix} \mathbf{d}_K \\ \mathbf{d}_F \end{pmatrix}, \text{ ahol } \mathbf{d}_K \in C^k, \mathbf{d}_F \in C^{n-1}.$$

A továbbiakban a hálózatok villamos állapotainak leírásával foglalkozunk.

DEFINÍCIÓ. Áramtérnek nevezzük és \mathcal{J} -vel jelöljük C^m következő alterét:

$$(F1.9) \quad \mathcal{J} = \{\mathbf{i} \mid \hat{\mathbf{Q}}\mathbf{i} = \mathbf{0}, \mathbf{i} \in C^m\}.$$

A definíció fizikai háttere *Kirchhoff I. törvényének* következő alakja: villamos hálózatokban az ág-áramok olyanok, hogy a hálózat bármely irányított vágására vonatkozóan előjeles összegük zérus.

Az F1.8., F1.11. tételekből következik, hogy C^m -nek az \mathcal{J} k -dimenziós altere, melynek ekvivalens megadási módjai az alábbiak:

$$(F1.10) \quad \mathcal{J} = \{\mathbf{i} \mid \mathbf{Q}\mathbf{i} = \mathbf{0}, \mathbf{i} \in C^m\},$$

$$(F1.11) \quad \mathcal{J} = \{\mathbf{i} \mid \hat{\mathbf{A}}\mathbf{i} = \mathbf{0}, \mathbf{i} \in C^m\},$$

$$(F1.12) \quad \mathcal{J} = \{\mathbf{i} \mid \mathbf{A}\mathbf{i} = \mathbf{0}, \mathbf{i} \in C^m\}.$$

Ezen összefüggések közül (F1.11) matematikai megfogalmazása *Kirchhoff I. törvénye* következő alakjának: Villamos hálózatok ág-áramai olyanok, hogy a hálózat bármely csomópontjára vonatkozóan az abba befolyó áramok összege egyenlő a kifolyó áramok összegével.

Az (F1.10) egyenlőségből (F1.8) felhasználásával adódik, hogy az áramteret definiáló egyenletrendszer a következő alakban is felírható:

$$(F1.13) \quad -\mathbf{F}^T \mathbf{i}_K + \mathbf{i}_F = \mathbf{0}.$$

A továbbiakban az $\mathbf{i} \in \mathcal{J}$ vektorok komponenseit ág-áramoknak is fogjuk nevezni.

Az (F1.13) összefüggésből látható, hogy a kötőélekre tetszőlegesen megválasztott ág-áramokhoz egyértelműen adódnak a fa-élek ág-áramai.

DEFINÍCIÓ. Feszültségtérnek nevezzük és \mathcal{V} -vel jelöljük C^m következő alterét:

$$(F1.14) \quad \mathcal{V} = \{\mathbf{v} \mid \hat{\mathbf{B}}\mathbf{v} = \mathbf{0}, \mathbf{v} \in C^m\}.$$

Fizikai hátterét tekintve a definíció *Kirchhoff II. törvényének* felel meg, mely szerint villamos hálózatokban az ág-feszültségek olyanok, hogy a hálózatban bármely irányított hurokra előjeles összegük zérus.

Az F1.8., F1.11. tételekből adódik, hogy \mathcal{V} dimenziója $n-1$, és \mathcal{V} ekvivalens módon a következőképp is megadható:

$$(F1.15) \quad \mathcal{V} = \{\mathbf{v} \mid \mathbf{B}\mathbf{v} = \mathbf{0}, \mathbf{v} \in C^m\}.$$

A feszültségtér definiáló egyenletrendszer (F1.8) felhasználásával a következő alakban is felírható:

$$(F1.16) \quad \mathbf{v}_K + \mathbf{F}\mathbf{v}_F = \mathbf{0}.$$

A $\mathbf{v} \in \mathcal{V}$ vektorok komponenseit ág-feszültségeknek is fogjuk nevezni.

Az (F1.16) összefüggésből leolvasható, hogy a fa-éleken tetszőlegesen megválasztva az ág-feszültségeket, a kötőélek ág-feszültségei egyértelműen adódnak. A feszültség és áramtér egymáshoz való viszonyát a következő tétel mutatja:

F1.12. TÉTEL. A \mathcal{V} feszültségtér és az \mathcal{J} áramtér egymás ortogonális kiegészítő alterei.

Bizonyítás. A definíciókból következően, az F1.10. tételt felhasználva $\mathcal{J} = \mathbf{C}_B$, $\mathcal{V} = \mathbf{C}_Q$ adódik, amiből az állítás az F1.10. tétel alapján következik.

A továbbiak során szükségünk lesz a következő jelölésre: Ha $\mathbf{a} \in \mathbf{C}^q$ ($q \geq 1$), akkor $\mathbf{a}^* \in \mathbf{C}^q$ jelölje azt a vektort, amely \mathbf{a} -ból úgy keletkezik, hogy annak komponenseit a komplex konjugáltakkal helyettesítjük.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ irányított gráfot, melynek éleihez a \mathcal{V} feszültségtér, illetve az \mathcal{J} áramtér pontjai alapján ág-feszültségeket, illetve ág-áramokat rendeltünk, hálózatként fogjuk nevezni.

DEFINÍCIÓ. Legyen $\mathbf{v} \in \mathcal{V}$ és $\mathbf{i} \in \mathcal{J}$. Ekkor a hálózat éleihez rendelt teljesítményt a következőképpen definiáljuk;

Ha $(j, l) \in \mathcal{A}^D$, akkor az s_{jl} -lél jelölt teljesítmény az alábbi:

$$(F1.17) \quad s_{jl} = v_{jl} i_{jl}^*.$$

Az élekhez rendelt teljesítmények vektorát jelölje \mathbf{s} , ennek komponenseit ág-teljesítményeknek is fogjuk nevezni.

A hálózathoz rendelt S összteljesítményt az ág-teljesítmények összegeként értelmezzük, így az F1.12. tétel felhasználásával kapjuk a következőt:

$$(F1.18) \quad S = \mathbf{v}^T \mathbf{i}^* = \langle \mathbf{v}, \mathbf{i} \rangle = 0.$$

Fizikai tartalmát tekintve \mathcal{J} és \mathcal{V} ortogonalitása teljesítmény megmaradást fejez ki.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ hálózat villamos állapotain az (\mathbf{i}, \mathbf{v}) , $\mathbf{i} \in \mathcal{J}$, $\mathbf{v} \in \mathcal{V}$ vektorpárokat értjük, így az összes villamos állapotok halmaza $\mathcal{J} \times \mathcal{V}$.

Fizikai hálózatokban általában nem minden (\mathbf{i}, \mathbf{v}) , $\mathbf{i} \in \mathcal{J}$, $\mathbf{v} \in \mathcal{V}$ pár megengedett, ezért bevezetjük a megengedett villamos állapotok fogalmát. Ez matematikai szempontból csak annyit jelent, hogy rögzítünk egy $\mathcal{M} \subseteq \mathcal{J} \times \mathcal{V}$ halmazt, és az (\mathbf{i}, \mathbf{v}) villamos állapotot megengedettnak nevezzük, ha $(\mathbf{i}, \mathbf{v}) \in \mathcal{M}$ teljesül.

A továbbiakban \mathcal{M} lehetséges rögzítéseire vonatkozóan két esetet tárgyalunk.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ hálózat (lineáris) admittancia-transzformációja a \mathbf{C}^m -et önmagába leképező olyan lineáris transzformáció, amely a \mathcal{V} feszültségtérre az \mathcal{J} áramtérbe képezi le. A transzformáció mátrixát a természetes bázisban jelölje \mathbf{Y} .

Ekkor az \mathbf{Y} mátrixnak a következő feltételt kell teljesítenie. Ha $\mathbf{v} \in \mathcal{V}$, akkor $\mathbf{Y}\mathbf{v} = \mathbf{i} \in \mathcal{J}$.

Jelölje \mathcal{J}_Y a \mathcal{V} feszültségtérnek a transzformáció által létrehozott képterét, más szóval:

$$(F1.19) \quad \mathcal{J}_Y = \{\mathbf{i} \mid \text{van olyan } \mathbf{v} \in \mathcal{V}, \text{ hogy } \mathbf{i} = \mathbf{Y}\mathbf{v} \text{ teljesül}\}.$$

Ekkor az \mathbf{Y} mátrixra vonatkozó feltétel megfogalmazható az $\mathcal{J}_Y \subseteq \mathcal{J}$ alakban is.

Amennyiben az $(\mathcal{N}, \mathcal{F}^D)$ hálózathoz hozzá van rendelve \mathbf{Y} mátrixú admittancia-transzformáció, akkor a megengedett villamos állapotok \mathcal{M}_Y halmaza $\mathcal{I}_Y \times \mathcal{V}$ következő részhalmaza lesz:

$$(F1.20) \quad \mathcal{M}_Y = \{(\mathbf{i}, \mathbf{v}) | \mathbf{i} = \mathbf{Y}\mathbf{v}, \mathbf{i} \in \mathcal{I}, \mathbf{v} \in \mathcal{V}\}.$$

Tekintsük most az $(\mathcal{N}, \mathcal{F}^D)$ feszítő fát és particionáljuk az \mathbf{Y} mátrixot a kötő és fa-éleknek megfelelően a következő módon:

$$(F1.21) \quad \mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{KK} & \mathbf{Y}_{KF} \\ \mathbf{Y}_{FK} & \mathbf{Y}_{FF} \end{pmatrix}.$$

Ekkor az admittancia transzformáció particionált alakban az alábbi lesz:

$$(F1.22) \quad \begin{pmatrix} \mathbf{i}_K \\ \mathbf{i}_F \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_{KK} & \mathbf{Y}_{KF} \\ \mathbf{Y}_{FK} & \mathbf{Y}_{FF} \end{pmatrix} \begin{pmatrix} \mathbf{v}_K \\ \mathbf{v}_F \end{pmatrix}.$$

Az (F1.16) alapján a \mathcal{V} feszültségtér előállítható a következő alakban:

$$(F1.23) \quad \mathcal{V} = \left\{ \mathbf{v} | \mathbf{v} = \begin{pmatrix} \mathbf{v}_K \\ \mathbf{v}_F \end{pmatrix} = \begin{pmatrix} -\mathbf{F} \\ \mathbf{E}_{n-1} \end{pmatrix} \mathbf{v}_F, \mathbf{v}_F \in \mathbb{C}^{n-1} \right\}.$$

Az (F1.22) összefüggésbe helyettesítve az alábbi adódik:

$$(F1.24) \quad \begin{pmatrix} \mathbf{i}_K \\ \mathbf{i}_F \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_{KK} & \mathbf{Y}_{KF} \\ \mathbf{Y}_{FK} & \mathbf{Y}_{FF} \end{pmatrix} \begin{pmatrix} -\mathbf{F} \\ \mathbf{E}_{n-1} \end{pmatrix} \mathbf{v}_F.$$

Innen a következőt kapjuk:

$$(F1.25) \quad \begin{aligned} \mathbf{i}_K &= (-\mathbf{Y}_{KK}\mathbf{F} + \mathbf{Y}_{KF})\mathbf{v}_F \\ \mathbf{i}_F &= (-\mathbf{Y}_{FK}\mathbf{F} + \mathbf{Y}_{FF})\mathbf{v}_F. \end{aligned}$$

A definíció alapján \mathbf{Y} pontosan akkor admittancia transzformáció, ha az (F1.25) alapján kapott $\mathbf{i} = \begin{pmatrix} \mathbf{i}_K \\ \mathbf{i}_F \end{pmatrix}$ eleme \mathcal{I} -nek, minden $\mathbf{v}_F \in \mathbb{C}^{n-1}$ esetén. Az áramteret definiáló (F1.13) összefüggés felhasználásával azt kapjuk, hogy minden $\mathbf{v}_F \in \mathbb{C}^{n-1}$ -re teljesülnie kell a következő összefüggésnek:

$$(F1.26) \quad (\mathbf{F}^T \mathbf{Y}_{KK} \mathbf{F} - \mathbf{F}^T \mathbf{Y}_{KF})\mathbf{v}_F = (\mathbf{Y}_{FK} \mathbf{F} - \mathbf{Y}_{FF})\mathbf{v}_F.$$

Innen az \mathbf{Y} részmátrixaira a következő feltétel adódik:

$$(F1.27) \quad \mathbf{F}^T \mathbf{Y}_{KK} \mathbf{F} - \mathbf{F}^T \mathbf{Y}_{KF} = \mathbf{Y}_{FK} \mathbf{F} - \mathbf{Y}_{FF}.$$

A továbbiakban azzal az esettel fogunk foglalkozni, amikor $\mathbf{Y}_{KF} = \mathbf{0}$, $\mathbf{Y}_{FK} = \mathbf{0}$. Ez, fizikai háttérét tekintve, azt jelenti, hogy csak olyan hálózatokat vizsgálunk, melyekben létezik olyan feszítő fa, hogy a fa-élek és a kötőélek közötti elektromágneses kapcsolat (kölcsonös admittanciák) elhanyagolható.

Ebben az esetben az admittancia-transzformáció mátrixa a következő lesz:

$$(F1.28) \quad \mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{KK} & \mathbf{0} \\ \mathbf{0} & -\mathbf{F}^T \mathbf{Y}_{KK} \mathbf{F} \end{pmatrix}.$$

Így az admittancia-transzformáció az alábbi alakot ölti:

$$(F1.29) \quad \mathbf{i}_K = \mathbf{Y}_{KK} \mathbf{v}_K,$$

$$(F1.30) \quad \mathbf{i}_F = \mathbf{F}^T \mathbf{Y}_{KK} \mathbf{F} \mathbf{v}_F.$$

Összefüggéseinkből (F1.16) felhasználásával az \mathcal{J}_Y altér következő előállítására adódik:

$$(F1.31) \quad \mathcal{J}_Y = \left\{ \mathbf{i} \mid \mathbf{i} = - \begin{pmatrix} \mathbf{Y}_{KK} \mathbf{F} \\ \mathbf{F}^T \mathbf{Y}_{KK} \mathbf{F} \end{pmatrix} \mathbf{v}_F, \mathbf{v}_F \in \mathbb{C}^{n-1} \right\}.$$

Az (F1.29), (F1.30) relációk fizikai háttere a következő. Az \mathbf{Y}_{KK} mátrix diagonális elemei a kötőelek ág-admittanciáinak, a mátrix többi eleme pedig a kötőelek közötti kölcsönös admittanciáknak felel meg. (A fizikai értelemben vett ág-admittanciákat a következő részben tárgyaljuk, lásd továbbá [69].) Figyelembe véve az (F1.16) összefüggést is, (F1.29), (F1.30) szerint: a faágak feszültségeit rögzítve, adott \mathbf{Y}_{KK} esetén egyértelműen meghatározottak az áramok a teljes hálózatban.

Megjegyezzük, hogy az admittancia-transzformáció \mathbf{Y} mátrixa természetesen nem azonos a hálózat passzív elemeinek (lásd [69]) admittanciáiból alkotott, a mérnöki szakirodalomban szintén \mathbf{Y} -nal jelölt mátrixszal.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}^D)$ hálózat (lineáris) impedancia-transzformációja a \mathbb{C}^m -et önmagába leképező olyan lineáris transzformáció, mely az \mathcal{J} áramteret a \mathcal{V} feszültségtérbe képezi le. A transzformáció mátrixát a természetes bázisban jelölje \mathbf{Z} . Ekkor a \mathbf{Z} mátrixnak a következő feltételt kell kielégítenie: Ha $\mathbf{i} \in \mathcal{J}$, akkor $\mathbf{Z}\mathbf{i} = \mathbf{v} \in \mathcal{V}$.

Jelöljük \mathcal{V}_Z -vel az \mathcal{J} áramtérnek a transzformáció által létrehozott képterét, azaz \mathbb{C}^m alábbi alterét:

$$(F1.32) \quad \mathcal{V}_Z = \{ \mathbf{v} \mid \text{van olyan } \mathbf{i} \in \mathcal{J}, \text{ melyre } \mathbf{v} = \mathbf{Z}\mathbf{i} \text{ teljesül} \}.$$

Ekkor a \mathbf{Z} mátrixra vonatkozó feltétel a következő alakban is megfogalmazható: $\mathcal{V}_Z \subseteq \mathcal{V}$.

Ha az $(\mathcal{N}, \mathcal{A}^D)$ hálózathoz hozzá van rendelve egy \mathbf{Z} mátrixú impedancia-transzformáció, akkor a megengedett villamos állapotok \mathcal{M}_Z halmaza az $\mathcal{J} \times \mathcal{V}_Z$ halmaz következő részhalmaza lesz:

$$(F1.33) \quad \mathcal{M}_Z = \{ (\mathbf{i}, \mathbf{v}) \mid \mathbf{v} = \mathbf{Z}\mathbf{i}, \mathbf{i} \in \mathcal{J}, \mathbf{v} \in \mathcal{V} \}.$$

Tekintsük az $(\mathcal{N}, \mathcal{A}^D)$ feszítő fát és particionáljuk a \mathbf{Z} mátrixot ennek megfelelően. Az admittancia-transzformációnál követett gondolatmenettel analóg módon eljárva, az alábbi \mathbf{Z} mátrix adódik:

$$(F1.34) \quad \mathbf{Z} = \begin{pmatrix} -\mathbf{F}\mathbf{Z}_{FF}\mathbf{F}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{FF} \end{pmatrix}.$$

Az impedancia-transzformáció a következő alakú lesz:

$$(F1.35) \quad \mathbf{v}_K = -\mathbf{F}\mathbf{Z}_{FF}\mathbf{F}^T \mathbf{i}_K$$

$$(F1.36) \quad \mathbf{v}_F = \mathbf{Z}_{FF} \mathbf{i}_F.$$

Innen (F1.13) felhasználásával a \mathcal{V}_Z altér alábbi előállítását kapjuk:

$$(F1.37) \quad \mathcal{V}_Z = \left\{ \mathbf{v} \mid \mathbf{v} = \begin{pmatrix} -\mathbf{F}\mathbf{Z}_{FF}\mathbf{F}^T \\ \mathbf{Z}_{FF}\mathbf{F}^T \end{pmatrix} \cdot \mathbf{i}_K, \mathbf{i}_K \in \mathbf{C}^k \right\}.$$

Az impedancia-transzformáció fizikai háttere a következő:

A \mathbf{Z}_{FF} mátrix diagonális elemei a fa-élek ág-impedanciáinak (lásd következő rész, illetve pl. [69]), a mátrix többi elemei pedig a faélek közötti kölcsönös impedanciáknak (lásd pl. [69]) felelnek meg. Tekintettel (F1.13)-ra is, az (F1.35) és (F1.36) összefüggések szerint: ha a kötőélek áramait rögzítjük, akkor adott \mathbf{Z}_{FF} esetén egyértelműen meghatározottak a feszültségek az egész hálózaton.

Természetesen az impedancia-transzformáció \mathbf{Z} mátrixa nem azonos a mérnöki szakirodalomban szintén \mathbf{Z} -vel jelölt, a passzív elemek impedanciáiból felépített mátrixszal.

Az admittancia-, ill. impedancia-transzformációk \mathbf{Y} , ill. \mathbf{Z} mátrixai lehetnek szingulárisak, még akkor is, ha \mathbf{Y}_{KK} , ill. \mathbf{Z}_{FF} nem szinguláris mátrixok. Számunkra a feszültségtérre, ill. az áramtérre korlátozott transzformációk invertálhatósága a lényeges.

Az admittancia-transzformáció esetében arról van szó, hogy az $\mathbf{i} = \mathbf{Y}\mathbf{v}$ összefüggéssel megadott transzformációt tekintjük, azonban csak $\mathbf{v} \in \mathcal{V}$ -re. Ez egy $\mathcal{V} \rightarrow \mathcal{I}_Y$ lineáris leképezés. Ahhoz, hogy kölcsönösen egyértelmű legyen, nyilván szükséges a $\dim \mathcal{V} = n - 1 = k = \dim \mathcal{I}$ feltétel teljesülése. Az (F1.31) előállításból látható, hogy a leképezés pontosan akkor kölcsönösen egyértelmű, ha az ott szereplő mátrix oszlopai lineárisan függetlenek. Amennyiben az inverz transzformáció létezik, azt korlátozott impedancia-transzformációnak nevezzük. Létezése azt jelenti, hogy minden $\mathbf{i} \in \mathcal{I}_Y$ -hoz pontosan egy $\mathbf{v} \in \mathcal{V}$ létezik, melyre (\mathbf{i}, \mathbf{v}) megengedett villamos állapot.

Hasonló megjegyzések érvényesek az impedancia-transzformációra és analóg módon definiáljuk a korlátozott admittancia-transzformációt.

Fizikai hátterét tekintve, az admittancia-, ill. impedancia-transzformációknak a hálózatokhoz való hozzárendelése az *Ohm törvénynek* a hálózatokra vonatkozó általánosításának tekinthető.

Megjegyezzük, hogy az $(\mathcal{N}, \mathcal{F}^D)$ feszítő fa rögzítésére vonatkozó kiinduló feltevésünknek a tárgyalásmód egyszerűsítésén túlmenően elvi szerepe nincs, a fentebb belátott tételek következményeképpen a feszültség-, ill. áramtérhez más-más feszítő fát választva, az az illető alterekben csupán báziscserét jelent.

Illusztrációként az egyszerű áramkört tekintjük. Ezen a következő hálózatot értjük:

$$\mathcal{N} = \{1, 2\}, \mathcal{A}^D = \{(1, 2), (2, 1)\}, n = m = 2, k = 1.$$

Az áramtér az $\mathcal{I} = \{(i_1, i_2) \mid i_1 = i_2\}$ altér, a feszültségtér pedig a $\mathcal{V} = \{(v_1, v_2) \mid v_1 = -v_2\}$ altér lesz, mindkettő egydimenziós altér \mathbf{C}^2 -nek. Esetünkben az impedancia és az admittancia-transzformációk megadása egyenértékű valamelyik ágra — legyen az pl. (2, 1) — vonatkozóan egy

$$(F1.38) \quad v_{21} = Z_{21} i_{21}$$

alakú összefüggés rögzítésével, ahol $Z_{21} \in \mathbf{C}$, konstans,

Fizikai tartalmát tekintve (F1.38) az *Ohm törvény* szokásos alakja. Az impedancia-transzformáció mátrixa az alábbi:

$$(F1.39) \quad \mathbf{Z} = \begin{pmatrix} -Z_{21} & 0 \\ 0 & Z_{21} \end{pmatrix}.$$

A korlátozott impedancia és admittancia-transzformációk léteznek, egymás inverzei, ily módon kölcsönösen egyértelmű leképezést létesítve \mathcal{I} és \mathcal{V} között.

DEFINÍCIÓ. Az (F1.38) összefüggésben szereplő Z_{21} mennyiséget a (2, 1) él ág-impedanciájának, az $Y_{21} = \frac{1}{Z_{21}}$ mennyiséget pedig az él ág-admittanciájának nevezzük.

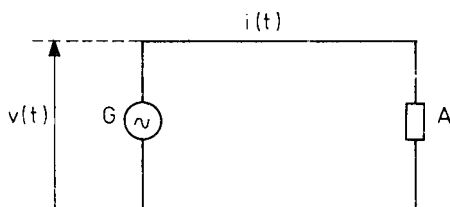
Végül megjegyezzük, hogy az ebben a részben tekintett elmélet \mathbf{C}^m helyett \mathbf{R}^m alapul vételével analóg módon tárgyalható.

F2. A villamosenergia-rendszer átviteli hálózatának fizikai leírása

Ebben a részben az átviteli hálózathoz egy helyettesítő, fizikai értelemben vett villamos hálózat konstrukcióját körvonalazzuk, mely helyettesítő hálózat feszültség-áram viszonyai a tényleges átviteli hálózat viszonyait az e részben ismertetett körülmények fennállása esetén kellő pontossággal tükrözik.

A harmonikus váltakozó áramokra vonatkozó előkészítő rész után az átviteli hálózat egy egyszerűsített leírása következik. Elsősorban a hálózat azon összetevőire koncentrálunk, melyek a feszültség-áram, és ezen keresztül a teljesítmény-viszonyokat lényegesen befolyásolják. A konstrukciót az átviteli hálózat főbb alkotóelemei fizikai értelemben vett helyettesítő villamos kapcsolásainak tárgyalásával folytatjuk. Végül megadjuk a helyettesítő villamos hálózat felépítését.

Az átviteli hálózat modelljének konstrukciója során harmonikus váltakozó áramokkal fogunk dolgozni, az ezekkel kapcsolatos fizikai fogalmak és alap-összefüggések felidézése céljából a következő egyszerű áramkört tekintjük:



F6. ábra

Az ábrán G feszültségforrás, A pedig a következő ideális hálózati elemek valamelyikét jelöli: ohmos ellenállás, induktivitás (tekercs), kondenzátor.

Ha a G feszültségforrás ω körfrekvenciájú, $v(t)$ harmonikus váltakozó feszültséget hoz létre, akkor az A helyére bármelyik fentebb felsorolt ideális hálózati elemet téve, az áramkörben ugyancsak ω körfrekvenciájú, $i(t)$ harmonikus váltakozó áram

folyik. Ez azt jelenti, hogy ha a feszültség

$$(F2.1) \quad v(t) = v_{\max} \cos(\omega t + \varphi),$$

akkor az áramkörben folyó áramra

$$(F2.2) \quad i(t) = i_{\max} \cos(\omega t + \psi)$$

érvényes, ahol v_{\max} , i_{\max} a feszültség, ill. áram csúcserkéi (amplitúdói), φ , ψ pedig a fázisszögek. Amennyiben A ohmos ellenállást jelöl R ellenállás-értékkel, akkor a feszültségre és áramra vonatkozóan minden időpillanatban a

$$(F2.3) \quad v(t) = Ri(t)$$

összefüggés érvényes.

Ha azonban A induktivitás, melynek L az indukció-együtthatója, akkor a feszültség és áram viszonyára a

$$(F2.4) \quad v(t) = L \frac{di(t)}{dt}$$

differenciál-összefüggés érvényes.

Hasonló a helyzet C kapacitású kondenzátor esetén, ahol a differenciál-összefüggés

$$(F2.5) \quad i(t) = C \frac{dv(t)}{dt}$$

alakú.

A következőkben a váltakozó áramokra vonatkozó komplex formalizmust vázoljuk. A dolgozat további részében jelölje j a komplex képzetes egységet.

Bevezetjük a következő jelöléseket:

$$(F2.6) \quad V = \frac{v_{\max}}{\sqrt{2}} e^{j\varphi},$$

$$(F2.7) \quad I = \frac{i_{\max}}{\sqrt{2}} e^{j\psi}.$$

Az (F2.6) egyenlőséggel bevezetett V -t, ill. az (F2.7) egyenlőséggel bevezetett I -t a feszültség, ill. áram fázis-reprezentációjának nevezzük, az $\frac{1}{\sqrt{2}}$ faktor szerepe a komplex teljesítmény tárgyalásánál válik majd világossá.

Az (F2.1) és (F2.2)-nek az (F2.3), (F2.4), (F2.5) összefüggésekbe való behelyettesítésével könnyen látható, hogy (F2.3), (F2.4), (F2.5) rendre formálisan ekvivalensek az alábbiakkal:

$$(F2.8) \quad V = RI$$

$$(F2.9) \quad V = j\omega LI$$

$$(F2.10) \quad V = \frac{1}{j\omega C} I.$$

A kapott formulák az *Ohm törvény* komplex reprezentációjának tekinthetők. Az összefüggésekben I együtthatójaként szereplő komplex mennyiségeket a megfelelő ideális áramköri elemek impedanciájának, azok reciprokát pedig admittanciájának nevezzük. Jelölje az ily módon meghatározott impedanciát Z , ennek értéke a különböző áramköri elemekre rendre az alábbi:

$$(F2.11) \quad Z = R \quad \text{ohmos ellenállásra,}$$

$$(F2.12) \quad Z = j\omega L \quad \text{induktivitásra,}$$

$$(F2.13) \quad Z = \frac{1}{j\omega C} \quad \text{kondenzátorra.}$$

A továbbiakban R -et rezisztív, ωL -et induktív, $\frac{1}{\omega C}$ -t pedig kapacitív reaktanciának nevezzük.

Amennyiben az eddig tekintett ideális áramköri elemekből és áramforrásokból felépítünk egy fizikai értelemben vett hálózatot, akkor az áram, ill. a feszültség pillanatnyi értékére fennállnak a *Kirchhoff-törvények*. Feltételezve a mindenütt azonos ω körfrekvenciát, könnyű belátni, hogy a *Kirchhoff-féle csomóponti törvény* ekvivalens módon megfogalmazható az áramok pillanatnyi értékeinek az áramok fázis-reprezentációjával való helyettesítése útján. Ugyanez a helyzet *Kirchhoff huroktörvénye* és a feszültségek fázis-reprezentációja vonatkozásában.

A továbbiakban feltételezzük valamennyi áram és feszültségforrással kapcsolatban az azonos, ω körfrekvenciát.

Látható, hogy az ideális áramköri elemekből felépített fizikai hálózat matematikai leírására alkalmazható az F1. részben leírt modell, a gráf éleinek egy-egy hálózati elemet megfelelően és fázis-reprezentációban gondolkodva. Egyszerűen belátható, hogy a hálózati elemeket a szokásos értelemben sorba, ill. párhuzamosan kötve a kapott összetett elemhez — a komplexen megfogalmazott *Ohm-törvényből* — hozzárendelhető impedancia-érték az egyenáramú esethez analóg módon kapható az alkotóelemek impedanciájából. Így a matematikai modell-hálózat méretének csökkentése céljából a fenti módon keletkező összetett elemnek csupán egy élt feleltetünk meg a gráfban.

Az F6. ábrán látható áramkörnek a matematikai értelemben vett (és az F1. részben definiált) egyszerű áramkör felel meg, a hálózati elemek impedanciájának pedig a gráf megfelelő éléhez rendelt ág-impedancia.

A harmonikus váltakozó áramokra vonatkozó teljesítmény-fogalom felidézése céljából most újra az F6. ábra áramkörét tekintjük. Ekkor a $p(t)$ -vel jelölt pillanatnyi teljesítmény a következő:

$$(F2.14) \quad p(t) = v(t)i(t) = v_{\max}i_{\max} \cos(\omega t + \varphi) \cos(\omega t + \psi).$$

Egyszerű átalakításokkal ez a következő alakra hozható:

$$(F2.15) \quad \begin{aligned} p(t) = & \frac{1}{2} v_{\max} i_{\max} \cos(\varphi - \psi) [1 + \cos 2(\omega t + \psi)] - \\ & - \frac{1}{2} v_{\max} i_{\max} \sin(\varphi - \psi) \sin 2(\omega t + \psi). \end{aligned}$$

Bevezetjük az alábbi jelöléseket:

$$(F2.16) \quad P = \frac{1}{2} v_{\max} i_{\max} \cos(\varphi - \psi),$$

$$(F2.17) \quad Q = \frac{1}{2} v_{\max} i_{\max} \sin(\varphi - \psi).$$

Jelöléseinkkel a pillanatnyi teljesítményre vonatkozó formula a következő alakot ölti:

$$(F2.18) \quad p(t) = P(1 + \cos 2(\omega t + \psi)) - Q \sin 2(\omega t + \psi).$$

P a hatásos teljesítmény, melyről könnyen megmutatható, hogy a pillanatnyi teljesítmény egy periódusra vett átlaga. Az (F2.18) első tagja egy, a P előjelétől függő irányú, a P középérték körül 2ω körfrekvenciával harmonikusan változó nagyságú teljesítmény-áramlásként interpretálható. ($P \geq 0$ a fogyasztásnak felel meg).

Q a meddő teljesítmény, az (F2.18) összefüggés második tagja szerint ez egy 0 átlagérték körüli, 2ω körfrekvenciájú harmonikus rezgés csúcserőve. Fizikailag a meddő teljesítmény egy, a feszültségforrás és az impedancia között végbemenő, 0 átlagértékű folyamatos energiacsere reprezentál. Pontosabban ez az energiacsere a forrás és az impedancia elektromágneses tere között megy végbe. Q fenti definíciójából és az (F2.4), (F2.5) egyenlőségekből az is egyszerűen levezethető, hogy az induktív impedancia meddő teljesítményt fogyaszt ($Q \geq 0$), míg a kapacitív impedancia meddőteljesítményt termel ($Q \leq 0$). A továbbiakban a meddő teljesítményt reaktív teljesítménynek is fogjuk nevezni.

A feszültség (F2.6), ill. áram (F2.7) fázis-reprezentációjának felhasználásával bevezetjük az S komplex teljesítményt a következő definícióval:

$$(F2.19) \quad S = VI^*.$$

A fázis reprezentációkat behelyettesítve, a komplex teljesítményre az alábbi alak kapható:

$$(F2.20) \quad S = \frac{1}{2} v_{\max} i_{\max} e^{j(\varphi - \psi)}.$$

Innen leolvasható, hogy S a hatásos és reaktív teljesítménnyel a következőképp fejezhető ki:

$$(F2.21) \quad S = P + jQ.$$

Látszólagos teljesítménynek nevezzük az $|S| = \sqrt{P^2 + Q^2}$ mennyiséget.

A dolgozatban a fizikai mennyiségeket a következő mértékegységekben tekintjük adottnak:

Impedanciák (rezisztív-, induktív-, kapacitív- reaktancia):	$[\Omega]$
Admittanciák (mindhárom):	$\left[S = \frac{1}{\Omega} \right]$
Feszültség:	[kV]
Áram:	[kA]
Hatásos teljesítmény:	[MW]
Meddő teljesítmény:	[Mvar]
Látszólagos teljesítmény:	[MVA]

Ezen előkészítés után megkezdjük a villamosenergia-rendszer átviteli hálózatának tárgyalását.

A villamosenergia-hálózatok általában többfázisúak, jellemző (és általunk vizsgálat alá vont) üzemállapotuk a szimmetrikus egyensúlyi állapot, azaz az egyes fázisokban az áram nagysága azonos, a fáziszögek egymáshoz képest konstans értékkel vannak eltolva. Ebben az esetben a hálózat villamos állapotának meghatározása egyfázisú helyettesítéssel végezhető el, a kapott eredményből a tényleges állapot rekonstruálható. Így a továbbiakban elegendő egyfázisú hálózatokkal foglalkoznunk. Az általunk vizsgált hálózatokban állandósult állapotban a harmonikus váltakozó áram frekvenciája mindenütt azonos. Így a modell felépítése során a fentiekben tárgyalt komplex fázis-reprezentációval dolgozhatunk.

A villamosenergia-rendszerben az erőművekben termelt villamos energiának a fogyasztókhoz való szállítása az átviteli hálózaton keresztül történik. Modellünkben ennek csak az ún. alaphálózat részét tekintjük, azaz a 750, 400, 220 kV-os ún. gerincvezetéseket és a 120 kV-os ún. főelosztóhálózatnak mérnöki szempontok alapján kiválasztott, az energiaelosztásban lényeges szerepet játszó részét.

A mérnöki szempontból tekintett hálózatban az energia szállítása távvezetéseken vagy kábeleken történik. A hálózat különböző feszültségintű részei transzformátorokon keresztül csatlakoznak egymáshoz.

A hálózat csomópontjai az erőművi gyűjtősínek és a fogyasztói alállomások. Az elnevezés onnan ered, hogy a hálózat ágai, az erőművek generátorai és transzformátorai és a fogyasztói kisebb feszültségű elosztóhálózatot az átviteli hálózathoz csatló transzformátorok rézből, vagy alumíniumból készült síneken keresztül kapcsolódnak egymáshoz egy-egy hálózati alállomáson.

Az alábbiakban röviden felsoroljuk azokat a főbb berendezéseket, amelyek együtteséből a csomópontokban teljesítménybetáplálás, ill. fogyasztás jelentkezik.

— Erőművi szinkrongenerátorok. Hatásos teljesítményt termelnek, a gerjesztésszabályozásuktól függően meddő teljesítményt termelhetnek vagy nyelhetnek. Modellünk szempontjából az egy csomópont-hoz csatlakozó összes generátor — esetleg több erőműben elhelyezve — együttes teljesítményét tekintjük.

— Fogyasztók (terhelés). Ezen az egy csomópont-hoz az alacsonyabb feszültség-szintű kiterjedt elosztóhálózaton keresztül csatlakozó összes fogyasztót értjük. A modellben az ily módon az egyes csomópontokban jelentkező együttes teljesítményigényt vesszük figyelembe.

Meddő teljesítmény betáplálására, ill. elnyelésére szolgálnak a következő eszközök:

— Szinkronkompenzátor. Szinkrongép, amely hatásos teljesítményt nem táplál a csomópont-ra. A forgórész gerjesztésétől függően akár meddő teljesítményt nyelőként, akár pedig kibocsátóként szerepeltethető, a szabályozási tartományból adódó korlátok között.

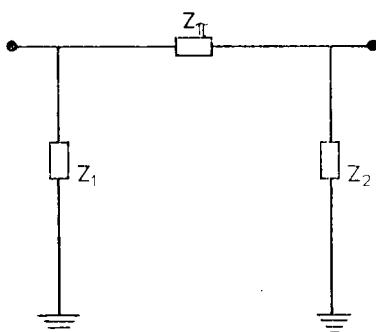
— Sönt-fojtótekeres. Meddő teljesítményt fogyasztó induktív berendezés.

— Sönt-kondenzátor. Meddő teljesítményt bocsát ki kapacitív jellege miatt. Egy csomópont-hoz több sönt-kondenzátor is csatlakozhat, ilyenkor a kapacitás adott határok között — diszkrét ugrásokkal — változtatható.

A következőkben az átviteli hálózat ágainak és transzformátorainak helyettesítő kapcsolásával foglalkozunk. Olyan egyszerű, néhány ágból álló, ideális hálózati ele-

mezből felépülő villamos kapcsolásról van szó, mely a benne szereplő impedancia-értékek alkalmas megválasztásával az átviteli hálózat megfelelő elemének feszültség-áram viszonyait a hálózat állandósult állapotában, mérnöki szempontból kellő pontossággal tükrözi. A helyettesítő kapcsolás természetesen csak az átviteli hálózat elemének stabil üzemi tartományában ad hű feszültség-áram képet.

A távvezetékek, kábelek, transzformátorok helyettesítésére egyaránt az F7. ábrán látható, 3 csomópontból és 3 ágból álló, ún. π -kapcsolás alkalmazható.



F7. ábra

A csomópontok közül az egyik a földpont, a másik kettő az átviteli hálózat gyűjtősíneinek felel meg. A Z_π , Z_1 , Z_2 impedancia-értékek meghatározása a távvezetékek, kábelek, ill. transzformátorok műszaki paramétereiből mérnöki megfontolások alapján történik [19]. Mindhárom esetben Z_π a következő alakú:

$Z_\pi = R + j\omega L$, ahol R a megfelelő elem ohmos ellenállásának, L pedig az indukтивitásnak felel meg. Az átviteli hálózatok fontos jellemzője, hogy valamennyi ágra az ωL induktív reaktancia lényegesen nagyobb az R ohmos ellenállásnál, transzformátorok esetében ezek a mennyiségek egy nagyságrendben különböznek.

Távvezetékek és kábelek esetében $Z_1 = Z_2 = \frac{2}{j\omega C}$, ahol C a távvezetékhez, ill. kábelhez rendelt kapacitás érték. Távvezetékek esetében az ωC érték kb. hat nagyságrenddel kisebb, mint ωL értéke. A kábelekhez rendelt kapacitás-érték lényegesen nagyobb a távvezeték kapacitás-értékénél.

A transzformátorok esetében a helyzet bonyolultabb, ezek egy része ugyanis változtatható áttételű, ún. szabályozható transzformátor. A nem szabályozható transzformátorokra, ill. a változtatható áttételű transzformátorok közül azokra, amelyek ún. középállásban vannak, az F7. ábrán látható helyettesítő kapcsolás egyszerűsödik, amennyiben a π „szárai” elhagyhatók, ezek tehát egy megfelelő Z_π impedanciájú ággal helyettesíthetők.

Azon szabályozható transzformátorok esetében, melyek nincsenek középállásban, az F7. ábrán látható Z_π , Z_1 , Z_2 ág-impedanciák értékei a transzformátor műszaki jellemzőiből és az áttételekből határozhatók meg [19].

A továbbiakban az egyszerűbb tárgyalásmód kedvéért feltesszük, hogy a szabályozható transzformátorok középállásban vannak, így valamennyi transzformátor egy-egy megfelelő Z_π impedanciájú ággal helyettesíthető.

Az eddigiek alapján felépíthetjük az átviteli hálózat fizikai modelljét, azaz egy olyan villamos hálózatot, melynek feszültség-áram viszonyai az átviteli hálózat feszültség-áram viszonyait a mérnöki alkalmazás szempontjából kellő pontossággal tükrözik.

A helyettesítő villamos hálózat csomópontjai a gyűjtősíneknek felelnek meg, jelölje ezek számát N . Feltételezzük továbbá, hogy a csomópontokat számokkal látjuk el. Bevezetjük $(N+1)$ -edik csomópontként a földpontot.

A hálózat ágai a következőképp adódnak. Tételezzük fel, hogy az i -edik és l -edik gyűjtősínt távvezeték vagy kábel köti össze. Az F7. ábrának megfelelő helyettesítő kapcsolásban az ellenállás, indukció-együttható és kapacitás értékeit jelölje rendre R_{il} , L_{il} , C_{il} . A helyettesítő villamos hálózatba a π felső szárának megfelel egy, az i -edik csomópontot az l -edikkel összekötő ág, $R_{il} + j\omega L_{il}$ ág-impedancia értékkel. A π szárai az i -edik, ill. l -edik csomópontot a földponttal kötik össze, mindkét ághoz a $\frac{2}{j\omega C_{il}}$ impedancia-érték van hozzárendelve.

Amennyiben az i -edik és l -edik gyűjtősínt transzformátor köti össze, és R_{il} , L_{il} ohmos ellenállást, ill. indukció-együtthatót rendeltünk hozzá, akkor a helyettesítő villamos hálózatban a transzformátornak egy, az i -edik és l -edik csomópontot összekötő ág felel meg, $R_{il} + j\omega L_{il}$ ág-impedancia-értékkel.

Ha az i -edik és l -edik csomópontot több távvezeték vagy kábel is összeköti, ezek a villamos hálózatban a párhuzamosan kötött impedanciákra érvényes helyettesítéssel a fenti három ágra redukálhatók, az egyszerűség kedvéért jelölje R_{il} , L_{il} , C_{il} az eredő ellenállást, indukció-együtthatót és kapacitást.

Jelölje M a fenti módon adódó villamos hálózatban azon ágak számát, amelyek mindkét végpontja gyűjtősínnek felel meg. Ezeket az egyszerűbb hivatkozás kedvéért a továbbiakban γ -típusú ágaknak fogjuk nevezni.

A vázolt eljárás eredményeként kapott villamos hálózatban a földponttól eltekintve minden csomóponthoz annyi kapacitív, másik végén földelt ág csatlakozik, amennyi az illető csomóponthoz illeszkedő γ -típusú ágak közül a távvezetéknek, ill. kábeleknak megfelelő ágak száma. Ezek a párhuzamosan kötött kapacitív ágak csomópontonként egyetlen ággal helyettesíthetők, melynek kapacitása az illető ágak kapacitásainak összege. Amennyiben valamely csomóponthoz csak transzformátoroknak megfelelő γ -típusú ágak csatlakoznak, akkor ahhoz nem vezetünk be ilyen ágat. Jelölje N_c azon csomópontok számát, melyekhez a fenti módon hozzárendeltünk az illető csomópontot a földponttal összekötő ágat, és tételezzük fel, hogy a csomópontok sorszámozása ezen csomópontokkal kezdődik. Ily módon az első N_c számú gyűjtősínnek megfelelő csomópontok mindegyikéhez egy-egy ilyen kapacitív ág csatlakozik, melyeket a továbbiakban β -típusú ágaknak fogunk nevezni. Jelölje C_i , $i=1, \dots, N_c$ ezen ágak kapacitásait.

Amint azt fentebb részletesebben tárgyaltuk, az átviteli hálózat gyűjtősínein több összetevő eredőjeként koncentrált teljesítmény betáplálás vagy igény jelentkezik. Ennek figyelembevétele céljából a síneknek megfelelő csomópontonként egy-egy további ággal egészítjük ki a villamos hálózatot, melyek az illető csomópontot a földponttal kötik össze. Úgy tekintjük, hogy a csomópontokon koncentráltan jelentkező teljesítmény betáplálás, ill. igény ezeken az ágakon keletkezik. Az így bevezetett N számú új ágat α -típusú ágaknak nevezzük a továbbiakban, melyekhez azokban ág-impedancia értéket nem tudunk hozzárendelni.

Az átviteli hálózatban a különböző távvezetékek, kábelek, transzformátorok elektromágneses kölcsönhatása elhanyagolható, ugyanezt feltételezzük tehát a villamos hálózat β - és γ -típusú ágairól. Feltételezhetjük azt is, hogy az α -típusú, valamint a β - és γ -típusú ágak között nincs elektromágneses kölcsönhatás. Ez azt jelenti, hogy a β - és γ -típusú ágakra vonatkozóan a feszültség-áram viszony ugyanolyan, mintha az illető ágak az F6. ábrán látható áramkörben az A -val jelölt áramköri elem helyén lennének.

Összefoglalva, a villamos hálózat a következő részekből áll. A csomópontok száma $N+1$, az $(N+1)$ -edik csomópont a földpont.

α -típusú ágak: az $1, \dots, N$ sorszámú csomópontokat a földponton összekötő ágak, ág-impedancia érték nincs hozzájuk rendelve. Ilyen élből N számú van a hálózatban.

β -típusú ágak: az $1, \dots, N_C$ sorszámú csomópontokat az $N+1$ sorszámú földponton kötik össze. Az i -edik ilyen ághoz $\frac{1}{j\omega C_i}$ ág-impedancia van hozzárendelve. A β -típusú ágak száma: N_C .

γ -típusú ágak: olyan ágak, melyek mindkét végpontjának sorszáma $\leq N$. Ezen ágak halmaza matematikailag az $1, \dots, N$ számokból alkotott összes számpárok halmazának részhalmazaként adható meg. Ha ez a számpár az $\{i, l\}$, akkor az ághoz rendelt ág-impedancia értéke: $R_{il} + j\omega L_{il}$. A γ -típusú ágak száma: M .

A hálózatnak az eddigiekben nem említett, fontos sajátosságai a következők: a földpontonól eltekintve, a csomópontoknak és a γ -típusú ágaknak megfelelő irányítatlan gráf összefüggő, továbbá a hálózat hurkolt, azaz a megfelelő gráf sok hurkot tartalmaz. Így feltételezhető, hogy $M > N$ teljesül.

F3. A villamosenergia-rendszer átviteli hálózatának matematikai modellje

E részben fő törekvésünk: a második részben megfogalmazott fizikai modellből kiindulva, az első részben adott általános elmélet alkalmazása révén az átviteli hálózat matematikai modelljének felépítése.

Először az átviteli hálózat matematikai modelljének alapját képező hálózatot építjük fel, majd annak feszültség- és áramterét vizsgáljuk, a hálózati veszteség fogalmát tárgyaljuk. Az admittancia-transzformáció analízise révén adódó egyenletrendszerből kiindulva bevezetjük a referenciapont fogalmát. E témák tárgyalása során mindkét megelőző fejezet jelöléseit használjuk, ahol ez konfliktushoz vezethet, esetenként tisztázzuk a jelöléseket. Végül megfogalmazzuk az átviteli hálózatokra vonatkozó hálózatszámítási feladatot, melyhez új, a téma műszaki szakirodalmához igazodó jelölésrendszert vezetünk be a villamos mennyiségekre vonatkozóan.

Az előző részben felépített villamos hálózatra alapozva, az átviteli hálózat matematikai modellje az F1. részben tárgyalt általános modellnek az alábbiakban megadott speciális esete lesz.

A hálózatot a következő $(\mathcal{N}, \mathcal{A}^D)$ gráffal reprezentáljuk:

\mathcal{N} : A gráf csúcsainak $n=N+1$ elemű halmaza, ahol a villamos hálózat földpontjának megfelelő csúcs az $N+1$ sorszámot kapja.

\mathcal{A}^D : A gráf éleinek $m = M + N + N_C$ elemű halmaza, melyet a villamos hálózat α -, β -, γ -típusú ágainak megfelelően particionálunk.

Legyen

$$(F3.1) \quad \mathcal{A}^D = \mathcal{A}_\alpha^D \cup \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D,$$

ahol az $\mathcal{A}_\alpha^D, \mathcal{A}_\beta^D, \mathcal{A}_\gamma^D$ él-halmazokat a következőképp definiáljuk:

\mathcal{A}_α^D : Ezek az élek a villamos hálózat α -típusú ágainak felelnek meg, halmazuk definíciója az alábbi:

$$(F3.2) \quad \mathcal{A}_\alpha^D = \{(1, N+1), (2, N+1), \dots, (N, N+1)\}.$$

\mathcal{A}_β^D : A villamos hálózat β -típusú ágának megfelelő élek halmaza, mely a következő lesz:

$$(F3.3) \quad \mathcal{A}_\beta^D = \{(N+1, 1), (N+1, 2), \dots, (N+1, N_C)\}.$$

\mathcal{A}_γ^D : A villamos hálózat γ -típusú ágainak megfelelő élek halmaza. Ebbe a halmazba M számú él tartozik, és ugyanazon sorszámú csúcsokat kötik össze, mint a megfelelő γ -típusú ágak a villamos hálózatban. Ezen élek irányítása a továbbiak szempontjából közömbös, mindenestre rögzítettnek tekintjük.

Az $(\mathcal{N}, \mathcal{A}^D)$ gráf éleit a következőképp sorszámozzuk: A sorszámozást az \mathcal{A}_α^D -ba tartozó éleken kezdjük, az (F3.2)-ben rögzített sorrendben, majd az \mathcal{A}_β^D -ba tartozó élek következnek az (F3.3)-ban megadott sorrendben, végül \mathcal{A}_γ^D elemei, tetszőleges sorrendben.

Jelölje \mathcal{N}_γ a továbbiakban a csúcsok halmazának a földpont elhagyásával adódó részhalmazát, azaz legyen $\mathcal{N}_\gamma = \{1, \dots, N\}$.

Az $(\mathcal{N}, \mathcal{A}^D)$ gráf triviálisan összefüggő, az előző rész végén tett megjegyzés értelmében feltételezzük, hogy az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ részgráf is összefüggő.

A feszítő fát az $\mathcal{F}^D = \mathcal{A}_\alpha^D$ előírással rögzítjük, vagyis feszítő fának az $(\mathcal{N}, \mathcal{A}^D)$ részgráfot választjuk. A fának $n-1 = N$ számú éle van. Az $(\mathcal{N}, \mathcal{A})$ gráfban az $(\mathcal{N}, \mathcal{A}_\alpha)$ részgráf ún. csillagot alkot, az \mathcal{A}_α halmazt az $N+1$ sorszámú csúcsot az összes többi csúccsal összekötő élek alkotják.

A feszítő fa fenti rögzítéséhez tartozó kötőélek halmazát $\mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D$ alkotja, a kötőélek száma: $k = m - n + 1 = M + N_C$. A kötőéleknek megfelelő $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ gráf az $(\mathcal{N}, \mathcal{A}^D)$ gráf összefüggő részgráfja.

Jelölje ebben a részben (az F1. résztől eltérően) $\hat{\mathbf{A}}$ az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ részgráf csúcs, él incidencia mátrixát, \mathbf{A} pedig az $\hat{\mathbf{A}}$ -ból az $(N+1)$ -edik sor törlésével keletkező redukált csúcs, él incidencia mátrixot.

Könnyen belátható, hogy a fa-vágások rendre az $1, \dots, N$ sorszámú csomópontokhoz tartozó csúcs-vágások lesznek az $(\mathcal{N}, \mathcal{A}^D)$ gráfra vonatkozóan, azaz sorra az $i = 1, \dots, N$ sorszámú csúcsokhoz illeszkedő, az \mathcal{A}^D halmazba eső élekből állnak. Így a redukált vágás-él incidencia-mátrix (F1.4) alapján esetünkben a következő alakot ölti:

$$(F3.4) \quad \mathbf{Q} = (\mathbf{A}, \mathbf{E}_N).$$

Ebből, valamint (F1.8)-ból következik, hogy a redukált hurok-él incidencia-mátrix megadható az alábbi formában, $\mathbf{F} = -\mathbf{A}^T$ helyettesítéssel:

$$(F3.5) \quad \mathbf{B} = (\mathbf{E}_{M+N_C}, -\mathbf{A}^T).$$

Az \mathcal{J} áramtér a fentiek szerint a \mathbf{C}^{M+N+N_c} tér $k=M+N_c$ dimenziós altere, definiáló egyenletrendszere (F3.4) és (F1.13) alapján a következő:

$$(F3.6) \quad \mathbf{A} \mathbf{i}_K + \mathbf{i}_F = 0.$$

Ebből az egyenletrendszerből látható, hogy amennyiben a kötőlélekre, azaz $\mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D$ elemeire vonatkozóan előírjuk az \mathbf{i}_K áram-vektort, akkor a fa-élek áramainak \mathbf{i}_F vektora egyértelműen adott.

Az (F3.6) összefüggés a következőképp is interpretálható: Ha az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ hálózatra szorítkozunk, akkor úgy is gondolkodhatunk, hogy annak földponttól különböző csúcsain árambetáplálás történik, a betáplált áramok rendre \mathbf{i}_F komponensei. Mivel A eme hálózat redukált csúcs, él incidenciamátrixa, ezért ebben a fel fogásban (F3.6) a csúcsokra betáplált és a hálózat élein folyó áramok közötti összefüggést fejez ki.

A \mathcal{V} feszültségtér az eddigiek alapján a \mathbf{C}^{M+N+N_c} tér $n-1=N$ dimenziós altere, melynek definiáló egyenletrendszere (F3.5) és (F1.16) alapján az alábbi:

$$(F3.7) \quad \mathbf{v}_K - \mathbf{A}^T \mathbf{v}_F = 0.$$

Az összefüggésből világos, hogy a fa éleire, azaz \mathcal{A}_β^D elemeire tetszőlegesen előírva a \mathbf{v}_F feszültség vektort, a kötőlélek ág-feszültségei egyértelműen adóttak. Ez az \mathcal{A}_β^D , ill. \mathcal{A}_γ^D halmazokba eső kötőlélekre vonatkozóan a következőket jelenti:

Ha a kötőél az \mathcal{A}_β halmazba esik, akkor $(N+1, i)$ alakú, ahol $1 \leq i \leq N_c$. Ebben az esetben az él által generált kötőhurok az $(N+1, i) \in \mathcal{A}_\beta^D$ és az $(i, N+1) \in \mathcal{A}_\alpha^D$ élekből áll. A gráf éleinek fentebb megadott sorszámozásának megfelelően a következőt kapjuk:

$$(F3.8) \quad v_{N+i} = v_i, \quad i = 1, \dots, N_c.$$

Ha az (i, l) kötőélre $(i, l) \in \mathcal{A}_\gamma^D$ teljesül, akkor a kötőhurok az $(i, l) \in \mathcal{A}_\gamma^D$, $(l, N+1) \in \mathcal{A}_\alpha^D$, $(i, N+1) \in \mathcal{A}_\alpha^D$ élekből áll. Így (F3.7) alapján a következő érvényes:

$$(F3.9) \quad v_{il} = v_i - v_l, \quad (i, l) \in \mathcal{A}_\gamma^D.$$

A kapott eredmények a következőképpen interpretálhatók: Rendeljünk az $(\mathcal{N}, \mathcal{A}^D)$ gráf első N csúcsához tetszőleges $u_i \in \mathbf{C}$, $(i=1, \dots, N)$ számértékeket, és legyen $u_{N+1}=0$, Legyen a $\mathbf{v} \in \mathbf{C}^{M+N+N_c}$ vektor az alábbi módon megadva.

Ha $(i, l) \in \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D$, akkor $v_{il} = u_i - u_l$, a fa-ágakra vonatkozóan pedig $v_i = u_i$, $i=1, \dots, N$.

Az (F3.8), (F3.9) összefüggésekből látható, hogy az így kapott \mathbf{v} vektor a feszültségtér eleme, azaz $\mathbf{v} \in \mathcal{V}$, továbbá az is világos, hogy a feszültségtér minden pontja megkapható ezen a módon. A feszültségtér pontjainak e módon történő előállítását az irodalomban „csomóponti potenciálok módszere” néven ismert [69].

A csúcsokhoz a fenti módon hozzárendelt u_i , $i=1, \dots, N+1$ számokat csomóponti potenciáloknak nevezzük, az ezek közül az első N számúból (azaz a földpontnak megfelelő csúcsot kihagyva) megalkotott $\mathbf{u} \in \mathbf{C}^N$ vektort a csomóponti potenciálok vektorának nevezzük. E vektort felhasználva (F3.7) a következőképp is írható:

$$(F3.10) \quad \mathbf{v}_K = \mathbf{A}^T \mathbf{u}.$$

A kapott eredmény interpretálható az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ részgráfra szorítkozva: a részgráf csúcsainak halmaza azonos az eredeti gráf csúcsainak halmazával, így a potenciálokat a részgráfra is vonatkoztathatjuk. Ekkor (F3.10) a részgráf ágfeszültségeinek a csomóponti potenciálokból történő szokásos potenciálkülönbségként való származtatását adja meg.

Az F2. részben tárgyaltak szerint a β - és γ -típusú ágak esetében, melyek a matematikai modell kötőeleinek felelnek meg, áganként érvényes az *Ohm-törvény*. Másszóval a komplex ágháramok és ágfeszültségek aránya konstans, függetlenül a többi ág komplex áramától és feszültségétől. A matematikai modellben ez azt jelenti, hogy érvényes az alábbi összefüggés:

$$(F3.11) \quad \mathbf{i}_K = \mathbf{Y}_{KK} \mathbf{v}_K,$$

ahol \mathbf{Y}_{KK} diagonális mátrix, melynek diagonális elemei a kötőelek ág-admittanciái. Így az F1. részben tárgyaltak szerint létezik az admittancia-transzformáció, melynek mátrixa (F1.28) (F3.5) alapján az alábbi:

$$(F3.12) \quad \mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{KK} & 0 \\ 0 & -\mathbf{A}\mathbf{Y}_{KK}\mathbf{A}^T \end{pmatrix},$$

ahol a szokásos módon, a kötő- és fa-éleknek megfelelően particionáltuk a mátrixot. Az admittancia-transzformáció a fa-élekre vonatkozóan a következő összefüggést adja:

$$(F3.13) \quad \mathbf{i}_F = -\mathbf{A}\mathbf{Y}_{KK}\mathbf{A}^T \mathbf{v}_F.$$

Bevezetjük a következő jelölést:

$$(F3.14) \quad \hat{\mathbf{Y}} = \mathbf{A}\mathbf{Y}_{KK}\mathbf{A}^T.$$

Így (F3.13) a következő alakot ölti:

$$(F3.15) \quad \mathbf{i}_F = -\hat{\mathbf{Y}} \mathbf{v}_F.$$

Az $\hat{\mathbf{Y}}$ mátrix csak az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ gráf struktúrájától és éleinek ág-admittancia értékétől függ. Az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ hálózatra szorítkozva, az irodalomban az $\hat{\mathbf{A}}\mathbf{Y}_{KK}\hat{\mathbf{A}}^T$ mátrixot csomóponti admittancia-mátrixnak nevezik (lásd pl. [69]). Az $\hat{\mathbf{Y}}$ mátrix ebből a földpontnak megfelelő sor és oszlop törlésével keletkezik.

Megjegyezzük, hogy ha az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ hálózatot lineáris, passzív $(N+1)$ -pólusként kezeljük (ennek definíciójára vonatkozóan lásd [69]), akkor (F3.15) alapján $-\hat{\mathbf{Y}}$ az ún. bemeneti admittancia-mátrix [69].

Az (F1.31) összefüggésből látható, hogy az admittancia-transzformáció a \mathcal{V} feszültséget az \mathcal{I} áramtér következő \mathcal{I}_Y alterére képezi le:

$$(F3.16) \quad \mathcal{I}_Y = \left\{ \mathbf{i} \mid \mathbf{i} = \begin{pmatrix} \mathbf{Y}_{KK}\mathbf{A}^T \\ -\mathbf{A}\mathbf{Y}_{KK}\mathbf{A}^T \end{pmatrix} \mathbf{v}_F, \quad \mathbf{v}_F \in \mathbb{C}^N \right\}.$$

Az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ gráf összefüggőségét felhasználva az F1.8 tételből következik, hogy \mathbf{A} sorai lineárisan függetlenek. Mivel \mathbf{Y}_{KK} nem szinguláris, ezért előző megjegyzésünk felhasználásával egyszerűen belátható, hogy az (F3.16) összefüggésben szereplő mátrix rangja N , annak oszlopoi lineárisan függetlenek.

Innen látható, hogy az \mathcal{S}_Y altér N -dimenziós és a \mathcal{V} feszültség-térre megszorított admittancia-transzformáció kölcsönösen egyértelmű megfeleltetést létesít \mathcal{V} és \mathcal{S}_Y között. Eredményünk az F1. részben bevezetett terminológia szerint a korlátozott impedancia-transzformáció létezését jelenti.

Mivel az \mathcal{S} áramtér dimenziójára az F2. részben tárgyaltak szerint $k = M + N_c > N$ teljesül, ezért \mathcal{S}_Y valódi altere \mathcal{S} -nek. Így a megengedett villamos állapotok halmazában valamennyi $\mathbf{v} \in \mathcal{V}$ előfordul valamely $\mathbf{i} \in \mathcal{S}$ -vel párosítva, és ez fordítva nem igaz.

A teljesítmény egyensúlyt kifejező (F1.18) ortogonalitási reláció a következő alakban is felírható:

$$(F3.17) \quad \mathbf{v}_K^T \mathbf{i}_K^* + \mathbf{v}_F^T \mathbf{i}_F^* = 0.$$

Az F1. részben tárgyaltak szerint ez az összefüggés az áramtér és a feszültségtér definícióiból (vagyis a *Kirchhoff-törvényekből*) következik, itt megadjuk egy további levezetését az (F3.7), (F3.11), (F3.13) relációkból. Az (F3.13) összefüggésben mindkét oldalon áttérve a komplex konjugáltakra, majd balról szorozva a \mathbf{v}_F^T vektorral, az alábbi adódik:

$$(F3.18) \quad \mathbf{v}_F^T \mathbf{A} \mathbf{Y}_{KK}^* \mathbf{A}^T \mathbf{v}_F^* + \mathbf{v}_F^T \mathbf{i}_F^* = 0.$$

Felhasználva az (F3.7) és (F3.11) összefüggéseket, innen közvetlenül adódik (F3.17). Az (F3.18) összefüggés a következőképp is írható:

$$(F3.19) \quad -\mathbf{v}_F \mathbf{i}_F^T = \mathbf{v}_K^T \mathbf{i}_K^*.$$

DEFINÍCIÓ. Az $S^v = \mathbf{v}_K^T \mathbf{i}_K^*$ mennyiséget hálózati veszteségnek nevezzük.

Fizikai háttérét tekintve S^v az átviteli hálózaton létrejövő összes komplex teljesítmény-veszteséget reprezentálja. Az (F3.19) egyensúlyi egyenletnek tekinthető, amennyiben a baloldalon szereplő mennyiség az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ hálózatra annak csúcsein betáplált teljesítményt, míg S^v az ágakon létrejövő veszteséget reprezentálja. A bal oldali negatív előjel abból adódik, hogy az \mathcal{A}_α -ba tartó fa-ágakat a földpont felé irányítottuk.

Az \mathbf{u} csomóponti potenciál vektort felhasználva, az eddigiekből könnyen adódnak a teljesítmény veszteségre az alábbi formulák:

$$(F3.20) \quad S^v = -\mathbf{u}^T \mathbf{i}_F^*,$$

$$(F3.21) \quad S^v = \mathbf{u}^T \hat{\mathbf{Y}}^* \mathbf{u}^*,$$

$$(F3.22) \quad S^v = \mathbf{u}^T \mathbf{A} \mathbf{Y}_{KK}^* \mathbf{A}^T \mathbf{u}^*.$$

Ezek után ismét az (F3.13) összefüggést tekintjük. A $\mathbf{v}_F = \mathbf{u}$ helyettesítéssel áttérve a csomóponti potenciálokra (F3.13) a következő alakot ölti:

$$(F3.23) \quad \mathbf{A} \mathbf{Y}_{KK} \mathbf{A}^T \mathbf{u} = -\mathbf{i}_F.$$

A kapott összefüggést most az \mathbf{u} komponenseire, mint ismeretlenekre vonatkozó egyenletrendszernek tekintjük, \mathbf{i}_F -et adottnak véve. Ekkor (F3.23) az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ hálózat csúcsein létrejövő potenciál-eloszlást írja le, adott csúcsponti árambetáplálások mellett. A következőkben az (F3.23) egyenletrendszer $\hat{\mathbf{Y}} = \mathbf{A} \mathbf{Y}_{KK} \mathbf{A}^T$, $N \times N$ -méretű együttható mátrixát részletesebb vizsgálatnak vetjük alá.

Az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ összefüggő gráf \mathbf{A} redukált csúcs, él incidenciamátrixa az alábbi szerkezetű:

$$(F3.24) \quad \mathbf{A} = (\mathbf{E}_c, \mathbf{A}_\gamma),$$

ahol $\mathbf{E}_c = -\begin{pmatrix} \mathbf{E}_{N_c} \\ \mathbf{0} \end{pmatrix}$, \mathbf{E}_{N_c} egy $N_c \times N_c$ méretű egységmátrix, \mathbf{A}_γ pedig az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ gráf csúcs, él incidenciamátrixa. Az F1.8. tétel szerint az \mathbf{A} mátrix sorvektorai lineárisan függetlenek.

Particionáljuk az \mathbf{Y}_{KK} mátrixot a β - és γ -típusú éleknek megfelelően:

$$(F3.25) \quad \mathbf{Y}_{KK} = \begin{pmatrix} \mathbf{Y}_\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_\gamma \end{pmatrix},$$

ahol $\mathbf{Y}_\beta, \mathbf{Y}_\gamma$ a megfelelő méretű diagonális mátrixok.

Így (F3.14) alapján az együtthatómátrix a következő alakot ölti:

$$(F3.26) \quad \hat{\mathbf{Y}} = \begin{pmatrix} \mathbf{Y}_\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \mathbf{A}_\gamma \mathbf{Y}_\gamma \mathbf{A}_\gamma^T.$$

Bevezetjük a következő jelöléseket:

$$(F3.27) \quad \mathbf{Y}_\beta = j\mathbf{Y}_\beta^K, \quad \mathbf{Y}_\gamma = \mathbf{Y}_\gamma^v + j\mathbf{Y}_\gamma^K,$$

ahol $\mathbf{Y}_\beta^K, \mathbf{Y}_\gamma^v, \mathbf{Y}_\gamma^K$ valós mátrixok. Ekkor egyszerű számolással adódik, hogy az (F3.23) komplex egyenletrendszernek megfelelő valós egyenletrendszer $\tilde{\mathbf{Y}}$ mátrixa az alábbi alakú:

$$(F3.28) \quad \tilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{0} & -\tilde{\mathbf{Y}}_\beta \\ \tilde{\mathbf{Y}}_\beta & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{A}_\gamma \mathbf{Y}_\gamma^v \mathbf{A}_\gamma^T & -\mathbf{A}_\gamma \mathbf{Y}_\gamma^K \mathbf{A}_\gamma^T \\ \mathbf{A}_\gamma \mathbf{Y}_\gamma^K \mathbf{A}_\gamma^T & \mathbf{A}_\gamma \mathbf{Y}_\gamma^v \mathbf{A}_\gamma^T \end{pmatrix},$$

ahol $\tilde{\mathbf{Y}}_\beta = \begin{pmatrix} \mathbf{Y}_\beta^K & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ és a $\mathbf{0}$ -k mindenütt a megfelelő méretű zérus-mátrixokat jelölik.

Az (F3.28) egyenlőségben az összeg első tagjaként szereplő mátrix elemei az F2. részben tárgyaltak szerint több nagyságrenddel kisebbek a második tag elemeinél, és a mátrix az $N_c < N$ esetben triviálisan szinguláris. A második tagként szereplő mátrixról az F1.8. tétel felhasználásával könnyű belátni, hogy az szinguláris. Mindezek alapján az egyenletrendszer mátrixa általában közel szinguláris (a fogalommal kapcsolatban lásd [57]), ami az egyenletrendszer numerikus megoldásánál komoly nehézségeket eredményez.

Ezek után az (F3.23) lineáris egyenletrendszert a következőképp módosítjuk: az \mathbf{u} potenciál-vektor valamely komponensét rögzítettnek tekintjük, és töröljük a rögzített komponensnek megfelelő sorszámú egyenletet. Az új egyenletrendszer mátrixát úgy kapjuk, hogy $\tilde{\mathbf{Y}}$ megfelelő sorszámú sorát és oszlopát töröljük. Ez a valós egyenletrendszer $\tilde{\mathbf{Y}}$ mátrixa két sorának és a megfelelő sorszámú két oszlopának törlését jelenti, és azt eredményezi, hogy (F3.28) második tagjában \mathbf{A}_γ szerepét az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ gráf megfelelő \mathbf{A}_γ^v redukált csúcs, él incidenciamátrixa veszi át. Az F1.8. tétel szerint \mathbf{A}_γ^v sorai lineárisan függetlenek, így mivel \mathbf{Y}_γ^v diagonális elemei pozitívak, ezért $\mathbf{A}_\gamma^v \mathbf{Y}_\gamma^v (\mathbf{A}_\gamma^v)^T$ könnyen láthatóan pozitív definit mátrix.

A módosított egyenletrendszer mátrixa, mind pedig annak az (F3.28) egyenlőséghez analóg felbontásában szereplő második tag-mátrixa nem szingulárisak, füg-

getlenül attól, hogy melyik sort töröltük. Ez a fentieket tekintetbe véve könnyen adódik a következő tételből:

F3.1. TÉTEL. Ha valamely \mathbf{D} , $2p \times 2p$ méretű mátrix a következő szerkezetű:

$$(F3.29) \quad \mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_3 \\ -\mathbf{D}_3 & \mathbf{D}_2 \end{pmatrix},$$

ahol \mathbf{D}_1 és \mathbf{D}_2 $p \times p$ méretű, szimmetrikus, pozitív definit mátrixok, \mathbf{D}_3 $p \times p$ méretű szimmetrikus mátrix, akkor \mathbf{D} nem szinguláris.

Bizonyítás: Legyen $\mathbf{x} \in \mathbb{R}^{2p}$ olyan vektor, melyre $\mathbf{D}\mathbf{x} = \mathbf{0}$ teljesül, és tekintsük az $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^p$ particiót. Feltételeinkből következően érvényes az alábbi:

$$(F3.30) \quad \mathbf{x}^T \mathbf{D} \mathbf{x} = \mathbf{x}_1^T \mathbf{D}_1 \mathbf{x}_1 + \mathbf{x}_2^T \mathbf{D}_2 \mathbf{x}_2 = 0.$$

Ez az összefüggés a \mathbf{D}_1 és \mathbf{D}_2 pozitív definit volta miatt implikálja, hogy $\mathbf{x} = \mathbf{0}$. Így \mathbf{D} oszlopai lineárisan függetlenek, azaz \mathbf{D} nem szinguláris, és a tételt bebizonyítottuk.

Mivel a fentiekben túlmenően a módosított egyenletrendszer mátrixára vonatkozó (F3.28) típusú felbontásban a második tag általában jól kondicionált mátrix, ezért tekintetbe véve a két tag-mátrix elemei közti nagyságrendi viszonyokat, a módosított egyenletrendszer mátrixa maga is jól kondicionált lesz.

DEFINÍCIÓ. Az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ hálózatnak azt a csúcsát, melyre vonatkozóan a komplex potenciált rögzítettnek tekintjük, a hálózat referenciapontjának nevezzük.

Fentebb kapott eredményünket a következő alakban is megfogalmazhatjuk: Amennyiben az $(\mathcal{N}, \mathcal{A}_\beta^p \cup \mathcal{A}_\gamma^p)$ hálózat referenciaponttól különböző csúcsaira előírjuk az áram-betáplálásokat, akkor az (F3.23) összefüggésből adódó módosított egyenletrendszer egyértelműen meghatározza a referenciaponttól különböző csúcsokra a komplex potenciálok értékeit, és az egyenletrendszer numerikus szempontból is megfelelő. Az (F3.23) összefüggésben a referenciapontnak megfelelő sorban az áram-betáplálás a potenciálokkal explicite ki van fejezve, ezért annak értéke a módosított egyenletrendszer megoldása után behelyettesítéssel megkapható. Az említett tulajdonságok függetlenek attól, hogy melyik csúcs tölti be a referenciapont szerepét.

A potenciál-eloszlás függ attól, hogy mely pontot választottuk referenciapontnak, és ezen belül természetesen függ a referenciapont komplex potenciáljától is.

Az átviteli hálózat referenciapontja nem választható meg tetszőlegesen. A fentiekből látható, hogy a referenciapontra vonatkozó áram-betáplálás nem írható elő, az a potenciál-eloszlásból egyértelműen adódik. Ezért referenciapontként csak azok a csúcsok szerepelhetnek, amelyeknél az áram-betáplálás szabályozható, azaz amelyekhez erőmű is csatlakozik, mégpedig elegendően nagy kapacitással ahhoz, hogy az igény szerinti betáplálást szolgáltatni tudja. Ezen túlmenően a referenciaponthoz csatlakozó erőmű általában fontos szerepet játszik a rendszer irányításában is (pl. frekvencia-szabályozás), melyekre e dolgozat keretei között nem térünk ki, az Olvasó további információt talál a [19] műben erre vonatkozóan.

A villamosenergia-rendszerek átviteli hálózatánál az áramok helyett a teljesítmény a ténylegesen figyelembe vett fizikai mennyiség, mind a termelés és fogyasztás, mind pedig a szállítás szempontjából. Így az (F3.23) összefüggést áramok helyett

teljesítményekkel kell megfogalmaznunk, ami a megfelelő egyenletrendszer linearitásának elvesztésével jár.

Jelölje $s \in \mathbb{C}^N$ a fa-ágakhoz rendelt teljesítmények vektorát, ekkor (F1.17) alapján s komponensei a következőképp kaphatók:

$$(F3.31) \quad s_p = u_p i_p^*, \quad i = 1, \dots, N.$$

Ezt és (F3.14) egyenlőségét felhasználva (F3.23) a következőképp írható:

$$(F3.32) \quad u_p^* \sum_{q=1}^N \hat{Y}_{pq} u_q = -s_p^*, \quad p = 1, \dots, N.$$

Az u potenciál-vektor komponenseit ismeretleneknek tekintve nemlineáris egyenletrendszer adódott. A továbbiakban ebből kiindulva megfogalmazzuk az átviteli hálózatokra vonatkozó hálózatszámítási ("load flow") feladatot.

Ehhez a villamos mennyiségekre vonatkozóan új jelölésrendszert fogunk bevezetni, ily módon igazodva a téma műszaki irodalmához. A függelék további részeiben már kizárólag ezt a jelölésrendszert használjuk.

A tárgyalásmód egyszerűsítése kedvéért a továbbiakban feltételezzük, hogy valamennyi csúcshoz csatlakozik β -típusú kapacitív ág, a $C_i = 0$, $i = N_c + 1, \dots, N$ előírásával.

A továbbiakban vizsgálatainkat az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ hálózatra fogjuk koncentrálni. Ennek N csúcsa és $N_c + M$ éle van. A csomóponti potenciálokkal fogunk dolgozni, jelölje ezek vektorát $U \in \mathbb{C}^N$.

Legyen $U_k = V_k e^{j\theta_k}$, $k = 1, \dots, N$, ahol V_k a potenciál abszolút értéke, θ_k a fázisszöge. A V_k , ill. θ_k , $k = 1, \dots, N$ komponensekből alkotott vektorokat jelölje V , ill. θ .

Szükségünk lesz a *Descartes-koordinátás* alakra is, itt kisbetűket fogunk a jelölésrendszerben alkalmazni, amivel az ütemezési modell jelölésrendszeréhez igazodunk.

Ennek megfelelően legyen $U = v + jw$, $v \in \mathbb{R}^N$, $w \in \mathbb{R}^N$.

A fentiekben megmutattuk, hogy a csomóponti potenciálok a modellben az α -típusú ágak ágfeszültségeiként interpretálhatók. A továbbiakban a csomóponti potenciálokat csomóponti feszültségeknek, ill. röviden feszültségeknek fogjuk nevezni. Fizikai tartalmukat tekintve ezek a síneknek a földponthoz viszonyított feszültségei, és a rendszer fontos, mért paraméterei.

Jelölje $S \in \mathbb{C}^N$ a fa-ágakhoz rendelt teljesítmények (-1) -szereseiből alkotott vektort, melynek komponensei a fentiek szerint az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ gráf csúcsaira történő teljesítmény-betáplálásként interpretálhatók.

Legyen $S = P + jQ$, $P \in \mathbb{R}^N$, $Q \in \mathbb{R}^N$; ahol P komponensei a hatásos, Q komponensei a meddő teljesítmények. A fogyasztást negatív betáplálásként értelmezzük.

Jelölje $I \in \mathbb{C}^M$ a hálózat γ -típusú éleihez tartozó ág-áramok vektorát és legyen $I = I^P + jI^Q$, ahol $I^P \in \mathbb{R}^M$ komponenseit valós áramoknak, míg $I^Q \in \mathbb{R}^M$ komponenseit képzetes vagy meddő áramoknak fogjuk nevezni.

A továbbiakban az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ éleinek olyan 1-től $N_c + M$ -ig terjedő sor-számozását tételezzük fel, mely a β -típusú élekkel kezdődik.

Az F2. részben tárgyaltak alapján a hálózat ág-admittanciái a következők lesznek:

$$(F3.33) \quad Y_{N+1,i} = j\omega C_i, \quad i = 1, \dots, N_c$$

és

$$(F3.34) \quad Y_{i,l} = \frac{1}{R_{il} + j\omega L_{il}}, \quad (i, l) \in \mathcal{A}_\gamma^D.$$

Bevezetjük az alábbi jelöléseket:

$$(F3.35) \quad \left. \begin{aligned} X_{il} &= \omega L_{il} \\ G_{il} &= \frac{R_{il}}{R_{il}^2 + X_{il}^2} \\ B_{il} &= \frac{X_{il}}{R_{il}^2 + X_{il}^2} \end{aligned} \right\} \quad (i, l) \in \mathcal{A}_\gamma^D.$$

Legyen $Y_{il} = Y_{il}$, $X_{il} = X_{il}$, $G_{il} = G_{il}$ és $B_{il} = B_{il}$, minden $(i, l) \in \mathcal{A}_\gamma^D$ esetén. Így a γ -típusú élek ág-admittanciáira az alábbi összefüggés adódik:

$$(F3.36) \quad Y_{il} = G_{il} - jB_{il}, \quad (i, l) \in \mathcal{A}_\gamma.$$

Az (F3.14) reláció jobb oldalán szereplő Y_{KK} mátrix diagonális elemei a fentebb megadott ág-admittanciák.

Könnyű belátni, hogy az (F3.14) relációval definiált \hat{Y} mátrix közvetlenül is megadható az alábbi formában:

$$(F3.37) \quad \hat{Y}_{ik} = \begin{cases} \sum_{l \in J(i)} Y_{il} + Y_{n+1,i}, & \text{ha } i = k, \\ -Y_{ik}, & \text{ha } i \neq k, k \in J(i), \\ 0, & \text{különben.} \end{cases}$$

Az (F3.37) előállításban és a továbbiak során is az F1. részben bevezetett $J(i)$ jelölés az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ gráfra vonatkozik.

Szükség lesz az \hat{Y} mátrix elemeinek valós és képzetes részeire is, így bevezetjük az alábbi jelölést:

$$(F3.38) \quad \hat{Y}_{ik} = \hat{G}_{ik} + j\hat{B}_{ik}; \quad \hat{G}_{ik} \text{ és } \hat{B}_{ik} \text{ valósak, } i, k = 1, \dots, N.$$

Kisebbszámú átalakítások után új jelölésrendszerünkben az (F3.32) relációk a következő alakot öltik:

$$(F3.39) \quad U_i \sum_{k=1}^N \hat{Y}_{ik}^* U_k^* = S_i, \quad i = 1, \dots, N.$$

Ebből egyszerű átrendezéssel, (F3.37) előállítást felhasználva, a következő összefüggést kapjuk:

$$(F3.40) \quad |U_i|^2 Y_{n+1,i}^* + \sum_{k \in J(i)} U_i (U_i - U_k)^* Y_{ik}^* = S_i, \quad i = 1, \dots, N.$$

Az (F3.39) egyenlőségeknek megfelelő valós összefüggések a feszültségek polár-koordinátás alakját véve, (F3.38) felhasználásával, a következők lesznek:

$$(F3.41) \quad \begin{aligned} V_i \sum_{k=1}^N V_k [\hat{G}_{ik} \cos(\Theta_i - \Theta_k) + \hat{B}_{ik} \sin(\Theta_i - \Theta_k)] &= P_i, \\ V_i \sum_{k=1}^N V_k [\hat{G}_{ik} \sin(\Theta_i - \Theta_k) - \hat{B}_{ik} \cos(\Theta_i - \Theta_k)] &= Q_i, \\ i &= 1, \dots, N. \end{aligned}$$

Az ekvivalens valós relációk *Descartes-koordináták* esetén:

$$(F3.42) \quad \begin{aligned} \sum_{k=1}^N (\hat{G}_{ik} [v_i v_k + w_i w_k] + \hat{B}_{ik} [v_k w_i - v_i w_k]) &= P_i, \\ \sum_{k=1}^N (\hat{G}_{ik} [v_k w_i - v_i w_k] - \hat{B}_{ik} [v_i v_k + w_i w_k]) &= Q_i, \\ i &= 1, \dots, N. \end{aligned}$$

Jelölje a továbbiakban $\Phi_i(\mathbf{V}, \boldsymbol{\Theta})$, ill. $\Psi_i(\mathbf{V}, \boldsymbol{\Theta})$ az (F3.41) egyenlőségekben a hatásos, ill. meddő teljesítménynél az összefüggések bal oldalán szereplő függvényt, hasonlóan definiáljuk az $f_i(\mathbf{v}, \mathbf{w})$, ill. $g_i(\mathbf{v}, \mathbf{w})$ függvényeket az (F3.42) relációkra vonatkozóan.

Ezekkel a jelölésekkel (F3.41) és (F3.42) az alábbi módon írhatók:

$$(F3.43) \quad \begin{aligned} \Phi_i(\mathbf{V}, \boldsymbol{\Theta}) &= P_i \\ \Psi_i(\mathbf{V}, \boldsymbol{\Theta}) &= Q_i \end{aligned} \quad i = 1, \dots, N,$$

$$(F3.44) \quad \begin{aligned} f_i(\mathbf{v}, \mathbf{w}) &= P_i \\ g_i(\mathbf{v}, \mathbf{w}) &= Q_i \end{aligned} \quad i = 1, \dots, N.$$

Az (F3.39) komplex, ill. az ekvivalens (F3.43) és (F3.44) valós összefüggések a csomóponti teljesítmény-betáplálások és a csomóponti potenciálok közötti kapcsolatot írják le. A továbbiakban a valós relációkat vizsgáljuk. Ezek nemlineáris összefüggések, a $2N$ teljesítmény típusú és a $2N$ feszültség típusú változót összesen $2N$ reláció köti össze. Csomópontonként 2 összefüggésünk és 4 változónk van. A változók közül csomópontonként kettőt lerögzítve és a másik kettőt ismeretlennek tekintve, az így nyert egyenletrendszerben az egyenletek száma meg fog egyezni az ismeretlenek számával.

Röviden diszkutálni fogjuk, hogy csomópontonként mely változók rögzíthetők le.

A hálózat referenciapontjánál a komplex feszültséget rögzítjük le, polár reprezentációban a fázisszöveget 0-nak véve. Ismeretlennek tekintjük a komplex teljesítmény-betáplálást. Az egyszerűség kedvéért feltételezzük, hogy a referenciapont 1-es sorszámmal.

A fogyasztói gyűjtősíneknek megfelelő csomópontokon adott a komplex teljesítmény-fogyasztás, ezt lerögzítjük, ismeretlennek véve a komplex feszültséget.

Ugyanez a helyzet azoknál a csomópontoknál, melyeknél a kapcsolódó erőművi generátorok rögzített meddőteljesítmény-betáplálásra vannak szabályozva.

A kétféle rögzítéshez képest közbülső esetnek tekinthetők azok a csomópontok, melyeken a hatásosteljesítmény-betáplálást és a feszültség abszolút értékét rögzítjük le. Ez a rögzítés feltételezi, hogy a csomópontokhoz csatlakozó berendezésekkel egyrészt az említett mennyiségek szabályozhatók, másrészt az (F3.43) alapján adódó meddő teljesítmény betáplálása megtörténik. Ilyenek azok a csomópontok, amelyekhez erőmű, ill. szinkronkompenzátor kapcsolódik. Az erőművek esetén a rögzítést indokolja, hogy az erőművi teljesítményszabályozással vezérelhető a hatásosteljesítmény-kibocsátás, a szinkrongenerátorok forgórészének áramával (gerjesztés) pedig a generátor kapocsfeszültsége és ezeken keresztül az erőművi gyűjtősín földhöz képesti feszültségének abszolút értéke. A szinkronkompenzátorral hasonló a helyzet, hiszen az is szinkrongép. Ezeket a csomópontokat feszültségtartó vagy meddőforrás csomópontoknak nevezzük és az egyszerűség kedvéért feltesszük, hogy sorszámaik: $2, \dots, L$.

Ezen kitérő után, a fentiek alapján adódik a csomópontok következő osztályozása és jelölése:

- a) Referenciapont: $V_1, \Theta_1=0$, rögzített; ismeretlenek: P_1 és Q_1 .
- b) (P, Q) -csomópontok: P_k és Q_k rögzített; ismeretlenek: V_k és Θ_k , $k=L+1, \dots, N$. Ilyen típusú a hálózati csomópontok zöme, 80–90%-uk.
- c) (P, V) -csomópontok: P_k és V_k rögzített; ismeretlenek: Q_k és Θ_k , $k=1, \dots, L$.

A hálózatszámítás alapfeladata ezek után a következőképp fogalmazható meg. Megoldandó az (F3.43), ill. az (F3.44) egyenletrendszer, azzal a kiegészítéssel, hogy a csomópontok típusától függően a megfelelő változók értékei rögzítettek. Az így kapott $2N$ egyenletből álló egyenletrendszerben $2N$ számú ismeretlen van.

Az egyenletrendszerekben a referenciapontnak, illetve a reaktív teljesítményre vonatkozó résznél a (P, V) -csomópontoknak megfelelő egyenleteket elhagyjuk, mert ezekben az egyik, csak abban az egyenletben szereplő ismeretlen a többivel az egyenletek által explicite ki van fejezve, így értéke a többi ismeretlen meghatározása után helyettesítéssel megkapható.

Megkaptuk a nagyfeszültségű villamosenergia-hálózatok hálózatszámítási feladatának egyenletrendszerét, az ún. „load flow” egyenleteket.

Jelölje V_k^0 , $k=1, \dots, L$ a lerögzített feszültségi abszolút értékeket.

A polár-koordinátás esetben a következő egyenletrendszer adódott:

$$\begin{aligned}
 \text{(F3.45)} \quad & P_i - \Phi_i(\mathbf{V}, \boldsymbol{\Theta}) = 0, & i = 2, \dots, N, \\
 & Q_i - \Psi_i(\mathbf{V}, \boldsymbol{\Theta}) = 0, & i = L+1, \dots, N, \\
 & V_k = V_k^0, & k = 1, \dots, L, \\
 & \Theta_1 = 0.
 \end{aligned}$$

Az egyenletrendszerben a rögzítések végrehajtása után ismeretlenek: V_i , $i=L+1, \dots, N$; Θ_i , $i=2, \dots, N$; az ismeretlenek száma: $2N-L-1$, ami megegyezik az egyenletek számával.

A Descartes-koordinátás változatnál adódik az alábbi egyenletrendszer:

$$\begin{aligned}
 (F3.46) \quad & P_i - f_i(\mathbf{v}, \mathbf{w}) = 0, & i = 2, \dots, N, \\
 & Q_i - g_i(\mathbf{v}, \mathbf{w}) = 0, & i = L+1, \dots, N, \\
 & v_k^2 + w_k^2 = (V_k^0)^2, & k = 1, \dots, L, \\
 & v_1 = V_1^0 \\
 & w_1 = 0.
 \end{aligned}$$

A rögzítések végrehajtása után az ismeretlenek: $v_i, i=2, \dots, N, w_i, i=2, \dots, N$, az ismeretlenek száma: $2N-2$, és ennyi egyenletünk van.

Megjegyezzük, hogy a meddőforrás-csomópontokon a meddőteljesítmény-kibocsátásra, ill. nyelésre vonatkozóan a kapcsolódó berendezések karakterisztikáitól függő korlátok érvényesek. Ha egy (P, V) csomóponton (F3.43) csak ezek megsértésével lenne kielégíthető, akkor azt (P, Q) csomóponttá kell átminősíteni. Hasonló a helyzet a (P, Q) csomópontok és a feszültség vonatkozásában. Ezt a problémát úgy szokás megoldani, hogy az (F3.43) hálózatszámítási feladatot megfogalmazzák a csomópontok egy rögzített minősítésével, és a megoldó algoritmust módosítják úgy, hogy az menet közben szükség szerint átminősíti a csomópontokat. E kérdés részleteivel jelen dolgozatunk keretei között nem foglalkozunk, lásd pl. [19].

F4. Teljesítményáramlás, Stott módszere a hálózatszámítási feladat megoldására

Modellünk az áramok helyett a csomópontokon betáplált teljesítményekre épül, ezért használni fogjuk a csomópontból az ágra „kiáramló” teljesítmény fogalmát. Ennek definiálása után a hálózatszámítási feladat megoldó algoritmusának konstrukciójának előkészítése céljából röviden összefoglaljuk a villamosenergia-rendszerek átviteli hálózatának számunkra fontos jellemző tulajdonságait. Ezután *Stott módszerének* ismertetése következik, végül a hálózati hatásosteljesítmény-áramlásra és veszteségre vonatkozó néhány összefüggést ismertetünk.

A teljesítményáramlás definíciójánál az (F3.40) összefüggésből indulunk ki. Bevezetjük az alábbi jelöléseket:

$$(F4.1) \quad S_{ik}^\gamma = U_i(U_i - U_k)^* Y_{ik}^*, \quad \text{ha } (i, k) \in \mathcal{A}_\gamma^D \quad \text{vagy} \quad (k, i) \in \mathcal{A}_\gamma^D,$$

$$(F4.2) \quad S_{n+1, k}^\beta = |U_i|^2 Y_{n+1, i}^*, \quad i = 1, \dots, N.$$

Jelöléseinkkel (F3.40) a következő alakot ölti:

$$(F4.3) \quad S_{n+1, i}^\beta + \sum_{k \in \mathcal{J}(i)} S_{ik}^\gamma = S_i, \quad i = 1, \dots, N.$$

Itt az S_i az i -edik csomópontba betáplált teljesítmény, és amennyiben az $S_{ik}^\gamma, S_{ik}^\beta$ mennyiségeket az i -edik csomópontból a csatlakozó ágakra kiáramló teljesítményeknek tekintjük, akkor (F4.3) formálisan *Kirchhoff csomóponti törvényének* analógonja teljesítményáramlásra. Könnyen látható, hogy fizikailag $S_{ik}^\beta, S_{ik}^\gamma$ valóban teljesítményt reprezentál.

Mindezek alapján az $(\mathcal{N}, \mathcal{A}_\beta^D \cup \mathcal{A}_\gamma^D)$ hálózatra vonatkozó teljesítményáramlásnak fogjuk nevezni az (F4.1), (F4.2) relációkkal definiált $S_{ik}^\beta, S_{ik}^\gamma$ mennyiségek összes-

ségét. A megfelelő valós, ill. képzetes részek összességét hatásos-, ill. reaktívteljesítmény-áramlásnak nevezzük.

Ily módon minden $(i, k) \in \mathcal{A}_Y^D$ ághoz két teljesítmény-értéket rendeltünk, melyek az ebben az összefüggésben irányítatlannak tekintett ág végpontjairól az ágra „ki-áramlanak”. Fontos megjegyzés, hogy $S_{ik}^Y \neq -S_{ki}^Y$, azaz az i -edik csomópontból kiáramló teljesítmény nem azonos a másik csomóponttra „befolyó” teljesítménnyel. Valóban, a k -edik csomóponttra befutó teljesítmény felfogható úgy, mint az ebből a csomópontból az (i, k) ágra kimenő teljesítmény -1 -szerese. Ily módon, ha a két teljesítmény különbségét S_{ik}^v -vel jelöljük, az alábbi adódik:

$$(F4.4) \quad S_{ik}^v = S_{ik}^Y + S_{ki}^Y = |U_i - U_k|^2 Y_{ik}^*, \quad (i, k) \in \mathcal{A}_Y^D.$$

A kapott S_{ik}^v mennyiségek fizikai tartalmukat tekintve az $(\mathcal{N}_Y, \mathcal{A}_Y^D)$ hálózat élein létrejövő teljesítmény-veszteségek.

Az F3. részben definiált S^v hálózati teljesítmény veszteségre a következő reláció adódik:

$$(F4.5) \quad S^v = \sum_{k=1}^N |U_k|^2 Y_{n+1,k}^* + \sum_{(i,k) \in \mathcal{A}_Y^D} |U_i - U_k|^2 Y_{ik}^*.$$

Összefüggésünk könnyen megkapható a (F3.22) egyenlőségből, (F3.14) és (F3.37) felhasználásával.

A következőkben az algoritmus-konstrukció előkészítéseként az átviteli hálózat néhány sajátosságát tekintjük.

A hálózat egyes ágai transzformátorokat reprezentálnak, az ilyen ágak végpontjai között számottevő potenciálkülönbség van. A hálózatszámítási feladat megoldását célzó algoritmusok stabilitása megkívánja a csomóponti feszültségek megfelelő normálását, amelynek ezen túlmenően számos más előnye is van [19]. Kétféle eljárás használatos.

Az elsőt ezek közül közös feszültségalapra redukálásnak nevezzük. Lerögzítünk egy feszültségértéket, majd valamennyi csomóponti feszültséget e körüli értékre normáljuk. A rögzített értéket nevezzük feszültségi konstansnak. A [100 kV, 200 kV), [200 kV, 400 kV), [400 kV, 700 kV), [700 kV, 900 kV) feszültség-intervallumok mindegyikéhez tartozik egy-egy névleges feszültségérték, a transzformáció ezek után a következő lesz:

$$(F4.6) \quad V_{új} = V_{rég} \frac{\text{feszültségi konstans}}{V_{névleges}}.$$

Ahhoz, hogy ugyanazok a teljesítmény-értékek adódjanak, mint a transzformáció előtt, természetesen a többi elektromos mennyiséget is ennek megfelelően transzformálnunk kell.

A második eljárás a viszonylagos egységek (v. e.) bevezetését jelenti. Ez leggyakrabban azt jelenti, hogy a csomóponti feszültségeket 1-körüli értékekre normáljuk. A transzformáció az alábbi:

$$(F4.7) \quad V_{v.e} = \frac{V_{rég}}{V_{névleges}}.$$

Ezzel az eljárással fizikai dimenzió nélküli értékek adódnak és az előző eljárásnál tett megjegyzésnek megfelelően itt is transzformálnunk kell a többi elektromos mennyiséget is.

Következő lépésként összefoglaljuk a villamosenergia-rendszerek átviteli hálózata azon jellemzőit, melyek a hálózatszámítási feladat megoldására irányuló algoritmusok konstrukciójánál fontos szerepet játszanak.

a) A nagyfeszültségű villamos hálózatoknál a távvezetékek, kábelek X_{ik} induktív reaktanciája általában sokkal nagyobb az R_{ik} ohmos ellenállásnál, és egy nagyságrenddel nagyobb az arány a transzformátoroknál. Ez azt eredményezi, hogy

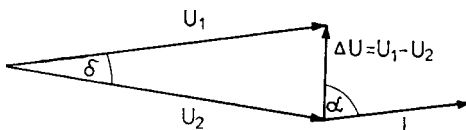
$$G_{ik} \ll B_{ik}, \quad \text{minden } (i, k) \in \mathcal{A}_\gamma^D\text{-re.}$$

b) A rendszer stabil működése során a referenciapontra vonatkoztatott Θ_k fázisszögek kicsik, $|\Theta_k| < 20^\circ$, $k = 1, \dots, N$. Az egyes ágakra vonatkozó fázis-különbségek is kicsik, $|\Theta_i - \Theta_k| < 10^\circ$, minden $(i, k) \in \mathcal{A}_\gamma^D$.

c) Normál üzemenet esetén a normált feszültségek eltérése kevesebb, mint 15%.

Tapasztalati tény, hogy a hatásosteljesítmény-áramlás és a feszültségeloszlás összefüggésében (melyet az (F3.43) egyenlőségrendszerben a P_i -re vonatkozó relációk írnak le) a feszültségek fázisszögei játsszák a domináns szerepet. Hasonló a tapasztalat a reaktívteljesítmény-áramlás és a feszültségek abszolút értékei vonatkozásában. A jelenség magyarázatához felvonultatunk néhány heurisztikus jellegű érvet [33].

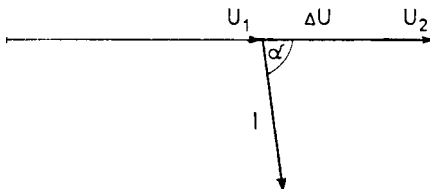
Tételezzük fel először, hogy az (1, 2) ág két végén adódó U_1 , ill. U_2 komplex potenciálra $|U_1| = |U_2|$ teljesül. Az ágon folyó áram legyen I .



F8. ábra

A fenti jellemzők szerint δ kicsi, továbbá $G_{12} \ll B_{12}$ miatt α közel 90° , így I és U_1 (ill. U_2) majdnem azonos fázisban vannak. Ezért a teljesítmény hatásos komponense nagy a reaktív komponenshez képest (abszolút értékben).

Ezután azt tételezzük fel, hogy a potenciálok azonos fázisban vannak, nagyságuk különböző.



F9. ábra

Ekkor I és U_1 (ill. U_2) fázisszögei közel 90° -kal különböznek, így a teljesítmény reaktív komponense nagy a hatásos komponenshez képest.

A jelenség további analízise céljából az (F3.41) egyenletrendszert felbontjuk a valós és képzetes részeknek megfelelő egyenletrendszerekre és egyszerű átalakításokat hajtunk végre. Bevezetjük a következő jelöléseket:

$$\begin{aligned} \delta_{ik} &= \Theta_i - \Theta_k, \\ (F4.8) \quad A(V_i, V_k, \delta_{ik}) &= V_i^2 - V_i V_k \cos \delta_{ik}, \\ B(V_i, V_k, \delta_{ik}) &= V_i V_k \sin \delta_{ik}, \end{aligned}$$

Ezekkel a jelölésekkel az egyenletrendszer a következő alakot ölti:

$$\begin{aligned} P_i &= \sum_{k \in J(i)} [G_{ik} A(V_i, V_k, \delta_{ik}) + B_{ik} B(V_i, V_k, \delta_{ik})], \\ (F4.9) \quad Q_i &= \sum_{k \in J(i)} [B_{ik} A(V_i, V_k, \delta_{ik}) - G_{ik} B(V_i, V_k, \delta_{ik})] - \omega C_i V_i^2, \\ i &= 1, \dots, N. \end{aligned}$$

Kiszámítjuk $A(V_i, V_k, \delta_{ik})$ és $B(V_i, V_k, \delta_{ik})$ parciális deriváltjait.

$$\begin{aligned} \frac{\partial A}{\partial V_i} &= 2V_i - V_k \cos \delta_{ik}, \quad \frac{\partial A}{\partial V_k} = -V_i \cos \delta_{ik}, \\ \frac{\partial A}{\partial \delta_{ik}} &= V_i V_k \sin \delta_{ik}, \\ (F4.10) \quad \frac{\partial B}{\partial V_i} &= V_k \sin \delta_{ik}, \quad \frac{\partial B}{\partial V_k} = V_i \sin \delta_{ik}, \\ \frac{\partial B}{\partial \delta_{ik}} &= V_i V_k \cos \delta_{ik}. \end{aligned}$$

Az előzőek szerint δ_{ik} kicsi, és ha a viszonylagos egységekre való normálást választjuk, akkor a feszültségek abszolút értéke 1-hez közeli. Ekkor $\left| \frac{\partial A}{\partial V_i} \right|$ és $\left| \frac{\partial A}{\partial V_k} \right|$ lényegesen nagyobbak (1-hez közeleiek), mint $\left| \frac{\partial A}{\partial \delta_{ik}} \right|$, (ami legfeljebb $\sin 10^\circ$ körüli érték lehet), azaz az $A(V_i, V_k, \delta_{ik})$ függvény elsősorban a potenciálok abszolút értékére érzékeny. Hasonlóan látható, hogy a $B(V_i, V_k, \delta_{ik})$ függvény elsősorban a fázisszög-különbségekre érzékeny.

Figyelembe véve, hogy $B_{ik} \gg G_{ik}$ minden $(i, k) \in \mathcal{A}_\gamma^D$ ágra, (F4.9)-ből adódik, hogy a hatásosteljesítmény-áramlás elsősorban a fázisszög-különbségekre, míg a reaktívteljesítmény-áramlás elsősorban a feszültségek abszolút értékére érzékeny.

Ezek után rátérünk az átviteli hálózat sajátosságait kihasználó (F3.45) hálózatszámítási feladat megoldására vonatkozó *Stott-módszer* ismertetésére [65]. Az algoritmus a *Newton-módszer* egy variánsa. A választást indokolja, hogy az átviteli hálózat fentebb tárgyalt sajátosságai következtében általában könnyű megadni elegendően jó induló megoldást.

Egy ilyen indulópont az ún. "flat voltage start", melynél valamennyi (P, Q) csomópont komplex feszültsége megegyezik a referenciapontéval, ugyanez igaz a (P, V) -csomópontoknál a fázisszögére. A *Newton-módszer* választását gyors konvergenciája is indokolja.

A teljesítményáramlás e részben vizsgált tulajdonságaira támaszkodva adódik *Newton módszerének* egy dekomponált változata.

Tételezzük fel, hogy adott egy $\mathbf{V}^{(0)}, \boldsymbol{\Theta}^{(0)}$ induló vektorpár, ahol $V_1^{(0)}, \dots, V_L^{(0)}$, az (F3.45)-ben előírt értékek. Ekkor, ha az $n+1$ -edik iteráció ($n \geq 0$) előtt adott $\mathbf{V}^{(n)}, \boldsymbol{\Theta}^{(n)}$ az algoritmus egy iterációja a következő lépésekből áll:

1. lépés: Végrehajtunk egy *Newton-korrekciós lépést*, $\boldsymbol{\Theta}^{(n)}$ -ből kiindulva, a következő egyenletrendszer megoldására:

$$(F4.11) \quad \begin{aligned} P_i - \Phi_i(\mathbf{V}^{(n)}, \boldsymbol{\Theta}) &= 0, \quad i = 2, \dots, N, \\ \Theta_1 &= \Theta_1^{(0)}, \end{aligned}$$

Az ily módon kapott vektor legyen $\boldsymbol{\Theta}^{(n+1)}$.

2. lépés: Végrehajtunk egy *Newton-korrekciós lépést* $V^{(n)}$ -et tekintve kiindulópontnak, a következő egyenletrendszerre:

$$(F4.12) \quad \begin{aligned} Q_i - \Psi_i(\mathbf{V}, \boldsymbol{\Theta}^{(n+1)}) &= 0, \quad i = L+1, \dots, N, \\ V_k &= V_k^{(0)}, \quad k = 1, \dots, L. \end{aligned}$$

Az így kapott vektor legyen $\mathbf{V}^{(n+1)}$.

Stott módszere [65] ebből az algoritmus-vázból kiindulva úgy konstruálható meg, hogy a *Newton-lépéseknél* alkalmazott *Jacobi-mátrixokban* a villamosenergia-hálózat sajátosságait kihasználva elhanyagolásokat hajtunk végre.

Az (F4.11) egyenletrendszerhez tartozó *Jacobi-mátrix* elemei:

$$(F4.13) \quad \begin{aligned} \left. \frac{\partial(P_i - \Phi_i(\mathbf{V}^{(n)}, \boldsymbol{\Theta}))}{\partial \Theta_k} \right|_{\boldsymbol{\Theta}=\boldsymbol{\Theta}^{(n)}} &= -V_i^{(n)} V_k^{(n)} [G_{ik} \sin(\Theta_i^{(n)} - \Theta_k^{(n)}) - \\ &\quad - B_{ik} \cos(\Theta_i^{(n)} - \Theta_k^{(n)})], \quad \text{ha } i \neq k, \\ \left. \frac{\partial(P_i - \Phi_i(\mathbf{V}^{(n)}, \boldsymbol{\Theta}))}{\partial \Theta_i} \right|_{\boldsymbol{\Theta}=\boldsymbol{\Theta}^{(n)}} &= \Psi_i(V^{(n)}, \boldsymbol{\Theta}^{(n)}) + B_{ii}(V_i^{(n)})^2, \\ &\quad i, k = 2, \dots, N. \end{aligned}$$

Az (F4.12) egyenletrendszer esetében a *Jacobi-mátrix* elemei:

$$(F4.14) \quad \begin{aligned} \left. \frac{\partial(Q_i - \Psi_i(\mathbf{V}, \boldsymbol{\Theta}^{(n+1)}))}{\partial V_k} \right|_{\mathbf{V}=\mathbf{V}^{(n)}} &= -V_i^{(n)} [G_{ik} \sin(\Theta_i^{(n+1)} - \Theta_k^{(n+1)}) - \\ &\quad - B_{ik} \cos(\Theta_i^{(n+1)} - \Theta_k^{(n+1)})], \quad \text{ha } i \neq k, \\ \left. \frac{\partial(Q_i - \Psi_i(\mathbf{V}, \boldsymbol{\Theta}^{(n+1)}))}{\partial V_i} \right|_{\mathbf{V}=\mathbf{V}^{(n)}} &= \frac{1}{V_i^{(n)}} [-\Psi_i(\mathbf{V}^{(n)}, \boldsymbol{\Theta}^{(n+1)}) + B_{ii}(V_i^{(n)})^2] \\ &\quad i, k = L+1, \dots, N. \end{aligned}$$

Az elhanyagolások alapját az e részben tárgyaltaiból következő alábbi relációk képezik:

$$(F4.15) \quad \cos(\Theta_i - \Theta_k) \sim 1, \quad |G_{ik} \sin(\Theta_i - \Theta_k)| \ll B_{ik},$$

minden $(i, k) \in \mathcal{A}_\gamma^D$ -ágra.

Ennek megfelelően az (F4.13), ill. (F4.14) relációkban a $G_{ik} \sin(\Theta_i - \Theta_k)$ tagokat elhanyagoljuk, és $\cos(\Theta_i - \Theta_k)$ -t 1-gyel helyettesítjük.

A módszer kifejlesztése során végzett számítógépes experimentáció eredményeképp az is kiderült, hogy a *Jacobi-mátrixban* $|\Psi_i|$ elhanyagolható a $B_{ii}V_i^2$ -hez képest.

Bevezetünk néhány jelölést. Legyen $\Delta\Theta = \Theta^{(n+1)} - \Theta^{(n)}$, ekkor persze $\Delta\Theta_1 = 0$ teljesül; $\Delta\mathbf{V} = \mathbf{V}^{(n+1)} - \mathbf{V}^{(n)}$, ekkor $\Delta V_i = 0$, $i = 1, \dots, L$.

Jelöléseinkkel a *Newton-korrekciókra* vonatkozó lineáris egyenletrendszerek az elhanyagolások és átrendezés után a következő alakot öltik:

$$(F4.16) \quad V_i^{(n)} \sum_{k=2}^N B_{ik} V_k^{(n)} \Delta\Theta_k = \Phi_i(\mathbf{V}^{(n)}, \Theta^{(n)}) - P_i, \quad i = 2, \dots, N,$$

$$(F4.17) \quad V_i^{(n)} \sum_{k=L+1}^N B_{ik} \Delta V_k = \Psi_i(\mathbf{V}^{(n)}, \Theta^{(n+1)}) - Q_i, \quad i = L+1, \dots, N.$$

Még egy elhanyagolással élünk, az (F4.16)-ban szereplő $V_k^{(n)}$, $k=2, \dots, N$ feszültség szintek értékét azonosnak vesszük. Ez a közös érték az F2. részben említett v. e. skálázás esetén 1, ha pedig a másik skálázást vesszük, akkor a feszültségi konstans (jelölve V_{SKAL}) értéke. Az utóbbi skálázást tekintve a korrekció egyenletrendszerei:

$$(F4.18) \quad \sum_{k=2}^N B_{ik} \Delta\Theta_k = \frac{\Phi_i(\mathbf{V}^{(n)}, \Theta^{(n)}) - P_i}{V_i^{(n)} V_{SKAL}}, \quad i = 2, \dots, N,$$

$$\sum_{k=L+1}^N B_{ik} \Delta V_k = \frac{\Psi_i(\mathbf{V}^{(n)}, \Theta^{(n+1)}) - Q_i}{V_i^{(n)}}, \quad i = L+1, \dots, N.$$

A megfelelő *Newton-lépések* ezek után a következők lesznek:

$$(F4.19) \quad \Theta^{(n+1)} = \Theta^{(n)} + \Delta\Theta, \quad \mathbf{V}^{(n+1)} = \mathbf{V}^{(n)} + \Delta\mathbf{V}.$$

A módszer nagy előnye, hogy az iterációk során a korrekciók kiszámításánál szereplő lineáris egyenletrendszerek mátrixa azonos. Így, ha a *CROUT-eliminációt* alkalmazzuk, a triangularizációt elegendő egyszer végrehajtani, a további iterációkban a megoldás már egyszerű visszahelyettesítéssel adódik.

A továbbiakban visszatérünk a teljesítményáramlás, illetve a hálózati veszteségek tárgyalására. A napi menetrend optimalizálásának modelljében szükségünk lesz a γ -típusú ágak esetében az (F4.1) egyenlőséggel definiált, az i -edik csomóponttól az (i, k) ágra áramló teljesítmény abszolút értékére, azaz a látszólagos teljesítményre. A csomóponti feszültségektől való függés kiemelése céljából jelölje ezt a teljesítményt $S_{ik}(v_i, v_k, w_i, w_k)$ az (F4.1) egyenlőségben szereplő felső indexet elhagyva az (F4.1) összefüggésből az alábbi alak adódik a látszólagos teljesítmény négyzetére:

$$(F4.20) \quad |S_{ik}(v_i, v_k, w_i, w_k)|^2 = (v_i^2 + w_i^2)[(v_i - v_k)^2 + (w_i - w_k)^2]|Y_{ik}|^2, \quad (i, k) \in \mathcal{A}_\gamma^D.$$

Itt a feszültségeket valós és képzetes részre bontva, azaz *Descartes-koordinátákkal* tekintettük, azt a szempontot szem előtt tartva, hogy a napi menetrend optimalizálására vonatkozó modellben a feszültségek valós és képzetes részei szerepelnek változóként. A Függelék befejező részében továbbra is ezt az alakot fogjuk használni.

Az $(\mathcal{N}_\gamma, \mathcal{A}_\gamma^D)$ hálózat élein létrejövő, az (F4.4) egyenlőségekkel megadott veszteséggel kapcsolatban bevezetjük a következő jelölést:

$$(F4.21) \quad S_{ik}^v = P_{ik}^v + jQ_{ik}^v, \quad P_{ik}, Q_{ik} \text{ valósak, } (i, k) \in \mathcal{A}_\gamma^D.$$

Innen a hatásosteljesítmény-veszteségre az alábbi adódik:

$$(F4.22) \quad P_{ik}^v = G_{ik} |U_i - U_k|^2.$$

A β -típusú ágak kapacitív jellege miatt nyilvánvaló, hogy azokon hatásosteljesítmény-veszteség nem jön létre.

Az (F4.22) összefüggést a távvezeték vagy transzformátort helyettesítő π -kapcsolás felhasználásával kaptuk. Pontosabb érték kapható, amennyiben — figyelembe véve, hogy a π „szárai” az ág felének kapacitív reaktanciáját helyettesítik — a kapacitás által termelt képzetes áramot a megfelelő ág-áram képzetes részéhez hozzáadjuk. Ezáltal tulajdonképpen a π -kapcsolással való helyettesítés révén keletkező, π elemen belüli hibát korrigáljuk. Az F2. részben a fizikai modell konstrukciója során az egyes távvezetéseket, kábeleket helyettesítő π -kapcsolások „szárait” csomópontként párhuzamosan kötött impedanciáknak tekintve egyetlen ággal helyettesítettük. A fentebb körvonalazott korrekcióhoz most szükségünk lesz az eredeti „száakra”. Jelölje C_{ik} az $(i, k) \in \mathcal{A}_\gamma^D$ ághoz az F2. részben tekintett helyettesítő kapcsolásban hozzárendelt kapacitás-értéket; amennyiben az ág transzformátort reprezentál, akkor legyen $C_{ik} = 0$. A π „szárai” a távvezeték, kábel kapacitásának felét reprezentálják, így a hozzájuk rendelt ág-admittancia $\frac{1}{2} j\omega C_{ik}$ lesz.

Ismeretes, hogy a hatásosteljesítmény-veszteségre érvényes a következő formula [19]:

$$(F4.23) \quad P_{ik}^v = R_{ik} |I_{ik}|^2.$$

A veszteség meghatározásához eszerint a komplex ág-áramot kell megadnunk. Alkalmazzuk az *Ohm-törvényt*:

$$(F4.24) \quad I_{ik} = \frac{U_i - U_k}{R_{ik} + jX_{ik}} + j\frac{1}{2} v_i \omega C_{ik}, \quad (i, k) \in \mathcal{A}_\gamma^D.$$

Az (F4.24) egyenlőségben a feszültségek *Descartes-koordinátás* alakját véve, az ág-áram az alábbi lesz:

$$(F4.25) \quad I_{ik} = (v_i - v_k) G_{ik} + (w_i - w_k) B_{ik} + j \left[(w_i - w_k) G_{ik} - (v_i - v_k) B_{ik} + \frac{1}{2} v_i \omega C_{ik} \right].$$

Innen a következő összefüggések olvashatók le:

$$(F4.26) \quad I_{ik}^P = -G_{ik}(v_k - v_i) - B_{ik}(w_k - w_i),$$

$$(F4.27) \quad I_{ki}^P = -I_{ik}^P,$$

$$(F4.28) \quad I_{ik}^Q = B_{ik}(v_k - v_i) - G_{ik}(w_k - w_i) + \frac{1}{2} v_i \omega C_{ik},$$

$$(F4.29) \quad I_{ki}^Q = -B_{ik}(v_k - v_i) + G_{ik}(w_k - w_i) + \frac{1}{2} v_k \omega C_{ik}.$$

Behelyettesítve (F4.23)-ban adódik a veszteség korrigált értéke:

$$(F4.30) \quad P_{ik}^v = \frac{1}{2} R_{ik} [2(I_{ik}^P)^2 + (I_{ik}^Q)^2 + (I_{ki}^Q)^2].$$

A továbbiakban az $(i, k) \in \mathcal{A}_\gamma^D$ ágon létrejövő hatásosteljesítmény-veszteséget a feszültségtől való függés kiemelése céljából $P_{ik}^v(v_i, v_k, w_i, w_k)$ -val jelöljük, minden $(i, k) \in \mathcal{A}_\gamma^D$ ágra.

Ennek megfelelően a hálózaton létrejövő összes veszteséget jelöljük $P^v(\mathbf{v}, \mathbf{w})$ -vel. Értéke az egyes ágakon létrejövő veszteségek összege:

$$(F4.31) \quad P^v(\mathbf{v}, \mathbf{w}) = \sum_{(i, k) \in \mathcal{A}_\gamma^D} P_{ik}^v(v_i, v_k, w_i, w_k).$$

Befejezésként néhány, a teljesítményáramlásra vonatkozó összefüggést tárgyalunk. Jelöljük $T_{ik}(v_i, v_k, w_i, w_k)$ -val az i -edik csomóponttól az $(i, k) \in \mathcal{A}_\gamma^D$ ágra áramló hatásos teljesítményt. Az (F4.1) egyenlőségből leolvasható, hogy ennek értéke az alábbi:

$$(F4.32) \quad T_{ik}(v_i, v_k, w_i, w_k) = G_{ik}[v_i(v_i - v_k) + w_i(w_i - w_k)] + B_{ik}[w_i v_k - w_k v_i].$$

Amennyiben az (i, k) ágra a kezdőpontból áramló hatásos teljesítményt meghatároztuk, akkor ebből és az ágon létrejövő veszteségből a k végpontból az ágra áramló teljesítmény meghatározható. Jelöljük ez utóbbit $T_{ki}(v_i, v_k, w_i, w_k)$ -val. A veszteség definíciójából adódik:

$$(F4.33) \quad T_{ki}(v_i, v_k, w_i, w_k) = -T_{ik}(v_i, v_k, w_i, w_k) + P_{ik}^v(v_i, v_k, w_i, w_k).$$

IRODALOM

- [1] BENDERS, J. F., "Partitioning procedures for solving mixed variables programming problems", *Numerische Mathematik* 4 (1962) 238—252.
- [2] BILLINGTON, R., SACHDEVA, M. S., "Real and reactive power optimization by suboptimum techniques", *IEEE Trans. PAS*—90 2 (1971) 950—956.
- [3] BÓKAY, B., RÁCZ, L., „Nagyfeszültségű hálózatok számítási módszerei”, *BME Továbbképző Intézet kiadványa*, Budapest, 1971.
- [4] CARPENTIER, J. W., "Differential injections method, a general method for secure and optimal load flows", in: *Proceedings of 8th PICA Conference*, (Minneapolis, 1973) 255—262.
- [5] CARPENTIER, J., SIROUX, J., "L'optimisation de la production à l'Electricité de France", *Bulletin de la Société Française des Electriciens* (1963).
- [6] COHEN, A. I., YOSHIMURA, M., "A branch-and-bound algorithm for unit commitment", *IEEE Trans. PAS*—102 2 (1983) 444—450.

- [7] COHN, N., *Control of Generation and Power Flow on Interconnected Power Systems* (John Wiley & Sons, New York, 1961).
- [8] CSIDA, S., NÉMETH, Á., „Villamosenergia-rendszerek teljesítményszabályozásának néhány kérdése”, *ERŐTERV Közlemények* 18 (1978) 53—63.
- [9] DANTZIG, G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, 1963).
- [10] DAY, J. T., „Forecasting minimum production costs with linear programming”, *IEEE Trans. PAS—90* 2 (1971) 814—823.
- [11] DEÁK, I., HOFFER, J., MAYER, J., PRÉKOPA, A., STRAZICKY, B., „A magyar villamosenergia-termelés optimális napi ütemezése” (modell és programleírás), Budapest, 1981.
- [12] DEÁK, I., HOFFER, J., MAYER, J., NÉMETH, Á., POTE CZ, B., PRÉKOPA, A., STRAZICKY, B., „A short description of the optimal daily scheduling of the electricity production in Hungary”, *MTA SzTAKI Working Paper MO/20*, 1980.
- [13] DEÁK, I., HOFFER, J., MAYER, J., NÉMETH, Á., POTE CZ, B., PRÉKOPA, A., STRAZICKY, B., „Optimal daily scheduling of the electricity production in Hungary”, in: *Large scale linear programming, Proceedings of a IASA Workshop*, editors: G. B. Dantzig, M. A. H. Dempster, Markku Kallio. IASA, Laxenburg, Austria, 1981, 923—960.
- [14] DEÁK, I., HOFFER, J., MAYER, J., NÉMETH, Á., POTE CZ, B., PRÉKOPA, A., STRAZICKY, B., „Optimal daily scheduling of the electricity production in Hungary”, *Operations Research in Progress*, editors: G. Feichtinger, P. Kall, (D. Reidel Publishing Company), 1982, 103—114.
- [15] DILLON, T. S., MARTIN, R. W., SJELOGREN, D., „Stochastic optimization and modelling of large hydrothermal systems for long-term regulation”, *IPC Business Press Electrical Power & Energy Systems* 2 1 (1980) 2—20.
- [16] DILLON, T. S., EDWIN, K. W., KOCHS, H. D., TEND, R. J., „Integer programming approach to the problem of optimal unit commitment with probabilistic reserve determination”, *IEEE Trans. PAS—97* 6 (1978) 2154—2166.
- [17] DOMMEL, H. W., TINNEY, W. F., „Optimal power flow solutions”, *IEEE Trans. PAS—87* (1968) 1866—1876.
- [18] DOPAZZO, J. F., KLITIN, O. A., STAGG, G. W., WATSON, M., „An optimization technique for real and reactive power allocation”, in: *Proceedings of the IEE*, 55 11 (1967) 1877—1885.
- [19] ELGERD, O. I., *Electrical Energy Systems Theory: An Introduction* (McGraw-Hill, New York, 1971).
- [20] ERISMAN, A. M., NEVES, K. W., DWARAKANATH, M. H., *Elektric Power Problems: The Mathematical Challenge* (SIAM, Philadelphia, 1980).
- [21] ESCUDERO, L. F., „Operation goals in maintenance scheduling for energy generators”, *IBM Working Paper*.
- [22] GESZTI, P. O., *Villamosművek* 1—2. (Tankönyvkiadó, Budapest, 1967).
- [23] GROSS, C. A., *Power System Analysis* (John Wiley & Sons, London, 1979).
- [24] HANO, I., TAMURA, Y., NARITA, S., MATSUMOMOTO, K., „Real time control of system voltage and reactive power”, *IEEE Trans. PAS—88* 10 (1969) 1544—1559.
- [25] HAPP, H. H., JOHNSON, R. C., WRIGHT, W. J., „Large scale hydro thermal unit commitment — method and results”, *IEEE Trans. PAS—90* 3 (1971) 1373—1384.
- [26] HAPP, H. H., „Optimal power dispatch — A comprehensive survey”, *IEEE Trans. PAS—96* (1977) 841—854.
- [27] HAPP, H. H., WIRGAU, K. A., „Static and dynamic VAR compensation in system planning”, *IEEE Trans. PAS—97* 5 (1978) 1564—1578.
- [28] HARHAMMER, P. G., „Economic dispatch using mixed integer programming”, in: *Proceedings of 5th Power System Computation Conference* (Cambridge, England, 1975) 1—17.
- [29] HARHAMMER, P. G., „Economic operation of multiple reservoir systems”, in: *Proceedings of 6th Power System Computation Conference* (Darmstadt 1978).
- [30] HARHAMMER, P. G., „Economic operation of electric power systems”, in: *Survey of Mathematical Programming*, ed. A. Prékopa, Akadémiai Kiadó, Budapest, 1979, Vol. 3., 179—194.
- [31] HOFFER, J., „Megengedett megoldások vizsgálatával bővített Benders dekompozíciós algoritmus”, *Alkalmazott Matematikai Lapok* 5 (1979) 177—190.
- [32] HOFFER, J., „Döntésektől függő ellátási feladatok megoldása Benders dekompozícióval”, *Alkalmazott Matematikai Lapok* 7 (1981) 73—82.
- [33] JOLISSAINT, CH. H., ARVANITIDES, N. V., LUENBERGER, D. G., „Decomposition of real and reactive power flows: A method suited to on-line applications”, *IEEE Trans. PAS—91* (1972) 661—670.

- [34] KELÉNYI, F., KISS, F., „Irodalmi áttekintés a KGST tagországokban alkalmazott operatív üzemelőkészítő programrendszerekről”, *VEIKI, kutatási jelentés*, 1979.
- [35] KÉRI, G., MOLNÁR, M., NÉMETH, Á., POTE CZ, B., PRÉKOPA, A., TURCHÁNYI, P., VIZVÁRI, B., „A magyar villamosenergia-rendszer napi menetrendjét meghatározó optimalizáló programozási modellek”, *MTA SZTAKI Operációkutatási Osztály*, Budapest, 1974.
- [36] KIRCHMAYER, L. K., *Economic Control of Interconnected Systems* (John Wiley & Sons, New York, 1959).
- [37] KIRCHMAYER, L. K., *Economic Operation of Power Systems* (John Wiley & Sons, New York, 1958).
- [38] KISS, L., POTE CZ, B., „Gazdaságossági optimalizációs eljárások villamosenergia-rendszerek irányításában”, *Ipargazdaság* 7 (1965).
- [39] KLOS, A., „Algebraic model of electrical network”, *IEEE Trans. PAS*—89 (1970) 240—262.
- [40] KOVÁCS, L. B., *A diszkrét programozás kombinatorikus módszerei* (Bolyai János Matematikai Társulat, Budapest, 1969).
- [41] KRUMM, L. A., »Obobsenie metoda niutona pri upravlении energetitskimi sistemami«, *Izvestiya Akademii Nauk USSR, Energetika i Transport* 3 (1976) 3—20.
- [42] LAMONT, J. W., ZESSO, W. G., RANTZ, M. W., „Daily fossil fuel management”, in: *Proceedings of PICA Conference*, 1979, 228—235.
- [43] LASDON, L. S., *Optimization for Large Systems* (The Macmillan Company, New York, 1972).
- [44] LAUER, G. S., SANDELL, N. R., BERTSEKAS D. P., POSBERGH, T. A., „Solution of large-scale optimal unit commitment problems”, *IEEE Trans. PAS*—101 1 (1982) 79—86.
- [45] LAUSANNE, B., VERET, C., »Description de la nouvelle version du modèle ORESTE de choix des démarrages des groupes thermiques«, *EDF—DER, Service Etudes de Réseaux HR* 32—0504.
- [46] LÉVAI, A., ZETTNER, T., *Hőerőművek IV.* (Műszaki Könyvkiadó, Budapest, 1971).
- [47] MVM—ERŐTERV—VEIKI—HITACHI, „Hungarian Automatic Load Dispatching (HALDIS), System Function Specification, Program Specification” (Budapest, 1979).
- [48] NÉMETH, Á., NAGY, I., „Daily load forecasting in the Hungarian National Control Center”—in: *Proceedings of VIIIth Power System Computation Conference* (Lausanne, 1981), 520—525.
- [49] MARIANI E., DI PERNA, A., „Programmazione giornaliera delle centrali idroelettriche a bacine e a serbatoio in un sistema di produzione misto”, *L' Energia Elettrica* 7 (1971) 427—448.
- [50] PANG, C. K., SHEBLE, G. B., ALBUEY, F., „Evaluation of dynamic programming based methods and multiple area representation for thermal unit commitments”, *IEEE Trans. PAS*—100 3 (1981) 1212—1218.
- [51] PESCHON, J., PIERCY, D. S., TINNEY, W. F., TVEIT, O. J., „Optimum control of reactive power flow”, *IEEE Trans. PAS*—87 (1968) 40—48.
- [52] POTE CZ, B., „Megszakításos üzemvitel rendszerszintű vizsgálata”, Műszaki doktori disszertáció, 1967.
- [53] POTE CZ, B., „Üzemállapot optimalizációs menetrend program”, „Számítástechnika a villamosenergia-iparban” konferencia kiadványa, Pécs, 1973.
- [54] POTE CZ, B., „Számológépes módszerek használata az energiarendszer operatív üzemelőkészítésben”, *VI. Országos Erőművi Konferencia kiadványa*, Budapest, 1974.
- [55] PRÉKOPA, A., *Lineáris programozás I.* (Bolyai János Matematikai Társulat, Budapest, 1968).
- [56] QUAZZA, G., „Highlights on technological trends in the on-line optimization of power system operation”, in: *Proceedings of IEE International Conference on On-line Operation and Optimization of Transmission and Distribution Systems* (London 22—25 June 1976).
- [57] RÓZSA, P., *Lineáris algebra és alkalmazásai* (Műszaki Könyvkiadó, Budapest, 1974).
- [58] SACHDEVA, M. S., BILLINGTON, R., PETERSON, C. A., „Representative bibliography on load forecasting”, *IEEE Trans. PAS*—96, 2 (1977) 697—700.
- [59] SASSON, A. M., ABOYTES, F., CARDENAS, K., GOMEZ, F., VILORIA, F., „A comparison of power systems static optimization techniques”, *ETZ—A Bd.* 93 (1972) 4. 9.
- [60] SESHU, S., REED, M. B., *Linear Graphs and Electrical Networks* (Addison—Wesley, London, 1961).
- [61] SGEF, *Système de gestion énergétique prévisionnelle*. Az Electricité de France, Service des Mouvements d'Energie kiadványa, Paris, 1981.
- [62] SHEN, C. M., LAUGHTON, M. A., „Power system load scheduling with security constraints using dual linear programming”, in: *Proceedings of IEE* Vol—117, 11 (1970) 2117—2127.
- [63] STAGG, G. W., EL-ABIAD, A. H., *Computer Methods in Power System Analysis*, (McGraw-Hill, New York, 1968).
- [64] STOTT, B., „Decoupled Newton load flow”, *IEEE Trans. PAS*—91 (1972) 1955—1959.
- [65] STOTT, B., ALSAÇ, O., „Fast decoupled load flow”, *IEEE Trans. PAS*—93 (1974) 859—869.

- [66] STOTT, B., "Review of Load Flow Calculation Methods", in: *Proceedings of IEEE* Vol—62, 7 (1974) 916—929.
- [67] SZENDY, K., *Korszerű hálózatszámítási módszerek*, (Akadémiai Kiadó, Budapest, 1967).
- [68] SZILÁGYI, L., „Villamos hálózatok optimális wattos teljesítményelosztásának modellje” (VEIKI, kutatási jelentés, Budapest, 1978).
- [69] VÁGÓ, I., *A gráfelmélet alkalmazása villamos hálózatok számításában*, (Műszaki Kiadó, Budapest, 1976).
- [70] VÁMOS, T., „Az energetikai optimalizálás néhány fő kérdéséről”, *VILLENKI Tanulmányok* 427, Budapest, 1962.
- [71] VÁMOS, T., „Erőművek számítógépes irányítása (A stockholmi IFAC/IFIP konferencia anyagából)”, *VILLENKI Tanulmányok*, Budapest, 1965.
- [72] WAIGHT, J. G., ALBUYEH, F., BOSE, A., "Scheduling of generation and reserve margin using dynamic and linear programming", *IEEE Trans. PAS*—100 5 (1981) 2226—2230.
- [73] WEEDY, B. M., *Electric Power Systems*, (John Wiley & Sons, London, 1967).

(Beérkezett: 1983. május 26.)

DEÁK ISTVÁN, HOFFER JÁNOS, MAYER JÁNOS,
PRÉKOPA ANDRÁS, STRAZICKY BEÁTA
MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1132 BUDAPEST, VICTOR HUGO U. 18.

NÉMETH ÁGOSTON
ERŐMŰ- ÉS HÁLÓZATTERVEZŐ VÁLLALAT
1054 BUDAPEST, SZÉCHENYI RKP.3.

POTECZ BÉLA
MAGYAR VILLAMOS MŰVEK TRÖSZT
ORSZÁGOS VILLAMOS TEHERELOSZTÓ
1014 BUDAPEST, PETERMAN BÍRÓ U. 5/7.

A LARGE SCALE, MIXED—VARIABLE MATHEMATICAL PROGRAMMING MODEL FOR THE SHORT—RANGE OPTIMAL SCHEDULE OF ELECTRIC POWER SYSTEMS WITH THERMAL POWER PLANTS TAKING NETWORK CONSTRAINTS INTO ACCOUNT

I. DEÁK, J. HOFFER, J. MAYER, Á. NÉMETH, B. POTECZ, A. PRÉKOPA, B. STRAZICKY

This unusually lengthy paper is the first detailed account of several years' work done at the Operations Research Department of the Computer & Automation Institute of the Hungarian Academy of Sciences dealing with the problem of daily scheduling of electricity production.

After the introduction giving a full account of the history of the work and Chapter 2 surveying the control tasks of electric power systems Chapter 3 contains a general mathematical programming model of a power generating system consisting of thermal power plants.

The model is general in the sense that for its formulation no assumptions were made to ensure numerical solvability of the nonlinear continuous-discrete problem corresponding to the model. Beside the conditions describing power plant relations the model contains the conditions necessary to the undisturbed operation of the transmission network.

Its objective function to be minimized is the total cost of power production which is the sum of the costs of fuel utilization of stillstand and restarting of power plant units and of network losses.

In Chapter 4 assumptions simplifying the scheduling problem, in Chapter 5 the model obtained on the basis of the above mentioned assumptions are set forth.

A large-scale mixed 0—1 linear programming problem with a special structure of the coefficient matrix corresponds to this simplified model, the solution of which is shown in Chapter 6.

A comprehensive appendix consisting of four parts summarizes the fundamentals and relations of the transmission network of electric power systems in a form easy to understand — in our opinion — by mathematicians and to serve as a basis for mathematical computations.

TRANZIENS KÉTFÁZISÚ FOLYADÉKÁRAMLÁS MATEMATIKAI MODELLEZÉSÉNEK PROBLÉMÁI

DÚS MAGDOLNA

Budapest

A cikk a tranziens kétfázisú áramlás két-közeg modelljeinek stabilitását vizsgálja. Foglalkozik a megoldásra használatos véges differencia módszer alkalmazhatóságának korlátaival, vizsgálja a modell pontosításának lehetőségeit.

1. Bevezetés

Az utóbbi évtizedben széleskörű kutatás folyik a tranziens (időben változó) egy- és kétfázisú áramlás minél pontosabb modellezésére. Ezek ismerete nagyon sok ipari, műszaki probléma vizsgálatához elengedhetetlen. Kutatásuk része az atomreaktorok biztonsági analizésének. Erre a célra egyre fejlettebb számítógépes programrendszerek kerülnek kifejlesztésre, melyek magja mindig egy áramlási modell a hűtőfolyadék áramlásának leírására. A biztonságos üzemeltetéshez és a lehetséges balesetekre való felkészüléshez fontos, hogy a megalkotott modellek minél pontosabban írják le a valóságot.

Az áramlásnak különböző fizikai modelljei ismeretesek. Ezek közül az utóbbi néhány évben kifejlesztett két-közeg modellek a legáltalánosabban használhatók. A modell lényege, hogy a két fázist különböző közegnek tekintik, melyek a rendelkezésre álló teret kitöltve együtt áramlanak úgy, hogy köztük sem termikus, sem mechanikus egyensúly nem áll fenn.

Az áramlást egy parciális differenciálegyenlet-rendszer írja le: a tömeg, az impulzus és az energia megmaradásának egyenletei. Ezek mellett szerepelnek segédegyenletekként a fázisok közötti kölcsönhatásokat leíró átadási tagok és a fázisok állapot-egyenletei, melyek az állapotváltozók közti függvénykapcsolatokat adják meg. A modell felírását a kezdeti és peremfeltételek megadása teszi teljessé.

Az érintett parciális differenciálegyenletek elmélete még nem teljesen kidolgozott. Problémák jelentkeznek a megoldásban, ha a fázisok sebességmezője különböző. Ilyenkor ugyanis a legáltalánosabban használt két-közeg modell egyenletrendszere nem teljesen hiperbolikus. Ezt a problémát vizsgáljuk a 2. pontban.

Az elsőrendű rendszernek van egy komplex konjugált karakterisztikus gyökpárja, emiatt a kezdeti érték feladat nem korrekt, a megoldás nem függ folytonosan a kezdeti értékektől. Tetszőleges kis perturbáció a kezdeti értékekben az időbeli megoldás során korlátlanul nőhet. Gyakorlati szempontból viszont a folytonos függés fontos lenne, mivel a kezdeti értékek általában mérés vagy közelítő számítás eredményei, tehát nem pontosak.

A megoldásra használt numerikus módszerek közül a véges differenciák módszere a legelterjedtebb. Egy ilyen módszert ismertetünk a 3. pontban. A véges differenciák alkalmazása azonban újabb problémákat vet fel. Mivel a feladat nem korrekt, ezért a vele konzisztens differencia sémák nem lesznek stabilak [6]. A 4. pontban azt vizs-

gáljuk a felírt konkrét numerikus sémára, hogy hogyan lehet mégis jó eredményeket kapni.

A valóságos áramlásokban fellép csillapítás (pl. a viszkozitás miatt). Az alap modell ezeket a jelenségeket elhanyagolja. Az 5. pontban azt vizsgáljuk, hogy ezeknek a figyelembe vétele hogyan befolyásolja a modell stabilitását.

2. A modell leírása

A két-közeg modellek általában 6 parciális differenciálegyenletből állnak, mindkét fázisra tartalmazzák a tömeg, az impulzus és az energia megmaradását leíró egyenleteket. Egyes modellek a két energia egyenlet helyett csak egyet, a keverék energia egyenletét foglalják magukba, a 6. egyenlet szerepét pedig valamilyen megszorítás játssza (pl. az a feltevés, hogy az éppen keletkező vagy eltűnő, tehát kisebb arányban jelenlevő folyadékfázis telítési hőmérsékleten van). Az alap egyenletrendszer mellett kezdeti és peremfeltételek adottak, valamint összefüggések a különböző átviteli tagokra. A modellezésnek fizikai szempontból ez a leginkább kérdéses része, mivel az átviteli tagok (tömeg, impulzus, energia átvadása a fázisok, illetve a tartály fala és a fázisok közt) még nem eléggé ismertek.

Az egyszerűség kedvéért a továbbiakban az egydimenziós két-közeg modell egyenletei közül a tömeg- és impulzus-megmaradási egyenleteket tekintjük. Ez a négy egyenlet tartalmazza a vizsgálat szempontjából lényeges jellemzőket. Az egyenletrendszer egyenes csőben történő izotermikus kétfázisú áramlást ír le.

Tömeg-megmaradás a gőzfázisra:

$$(2.1) \quad \frac{\partial}{\partial t}(\alpha \varrho_v) + \frac{\partial}{\partial x}(\alpha \varrho_v v_v) = 0.$$

Tömeg-megmaradás a folyadék fázisra:

$$(2.2) \quad \frac{\partial}{\partial t}[(1-\alpha)\varrho_l] + \frac{\partial}{\partial x}[(1-\alpha)\varrho_l v_l] = 0.$$

Impulzus-megmaradás a gőzfázisra:

$$(2.3) \quad \alpha \varrho_v \frac{\partial v_v}{\partial t} + \alpha \varrho_v v_v \frac{\partial v_v}{\partial x} + \alpha \frac{\partial P}{\partial x} = K(v_l - v_v).$$

Impulzus-megmaradás a folyadék fázisra:

$$(2.4) \quad (1-\alpha)\varrho_l \frac{\partial v_l}{\partial t} + (1-\alpha)\varrho_l v_l \frac{\partial v_l}{\partial x} + (1-\alpha) \frac{\partial P}{\partial x} = K(v_v - v_l).$$

Jelölések: α : gőz térfogattört,
 ϱ_v : gőz sűrűség,
 ϱ_l : folyadék sűrűség,
 v_v : gőz áramlási sebesség,
 v_l : folyadék áramlási sebesség,
 P : nyomás,
 K : impulzus átvadási tényező.

Az egyenletek felírása azzal a feltételezéssel történt, hogy az egyes fázisok különálló folyadékként viselkednek, közös nyomásmező hat rájuk és kitöltik a rendelkezésre álló térfogatot. Az egyenletek bal oldalainak első tagjai a tömeg, illetve az impulzus időbeli változását adják meg egy adott pontban. A második tagok a tömeg, illetve impulzus konvekcióját jelentik ugyanabból a pontból. Az impulzus egyenletekben a harmadik tagok azokat a nyomásból származó erőket foglalják magukba, melyek a fázisokat gyorsítják. Az impulzus egyenletek jobb oldalai a fázisok közti impulzus átadást fejezik ki. Ezek egyenlő nagyságúak, de ellenkező előjelűek.

Feltesszük, hogy a folyadék fázis összenyomhatatlan a gőzfázishoz képest. Ekkor a (2.1)–(2.4) egyenletrendszer μ karakterisztikus gyökeinek az alábbi egyenletet kell kielégíteniük:

$$(2.5) \quad \alpha Q_I (\mu - v_I)^2 + (1 - \alpha) Q_v (\mu - v_v)^2 - \alpha Q_I c_v^{-2} (\mu - v_v)^2 (\mu - v_I)^2 = 0,$$

ahol

$$(2.6) \quad c_v^{-2} = \frac{\partial Q_v}{\partial P}.$$

(c_v a hang terjedési sebessége a gőzfázisban.)

Beszorozva c_v^2 -tel és átrendezve azt kapjuk, hogy két valós gyök van, az egyik kevéssel kisebb, mint $v_v - c_v$, a másik kicsit nagyobb, mint $v_v + c_v$.

A polinom vizsgálata azt mutatja, hogy ha lenne még valós gyök, akkor annak ezek közé kellene esnie. Ez azonban $v_v \neq v_I$ esetén nem lehetséges, tehát a másik két gyök komplex, méghozzá egymás konjugáltjai. Ha $v_v, v_I \ll c_v$ (a gyakorlatban ez a jellemző), akkor a komplex gyökök értéke közelítőleg

$$(2.7) \quad \mu \cong \frac{v_I + \varepsilon^2 v_v}{1 + \varepsilon^2} \pm i \frac{\varepsilon}{1 + \varepsilon^2} (v_v - v_I),$$

ahol

$$\varepsilon^2 = \frac{(1 - \alpha) Q_v}{\alpha Q_I}.$$

Azt kapjuk eredményül, hogy az alapvető két-közeg modellt leíró egyenletrendszer nem tisztán hiperbolikus, a karakterisztikus gyökök közt komplexek is vannak, ha $v_v \neq v_I$, azaz az egyes fázisok különböző sebességgel mozognak.

Ennek az a következménye, hogy a megoldás nem függ folytonosan a kezdeti értékektől. A gyakorlatban pedig a folytonos függésnek nagy a jelentősége, hiszen az adatok általában mérések eredményei.

2.1. TÉTEL. Tekintsük a következő n egyenletből álló elsőrendű kvázilineáris parciális differenciálegyenlet-rendszert

$$(2.8) \quad \mathbf{A}(\mathbf{v}) \frac{\partial \mathbf{v}}{\partial t} + \mathbf{B}(\mathbf{v}) \frac{\partial \mathbf{v}}{\partial x} + \mathbf{c}(\mathbf{v}) = \mathbf{0},$$

ahol $\mathbf{v} = (v_1, v_2, \dots, v_n)^T$ a függő változók n dimenziós oszlopvektora, \mathbf{A} és \mathbf{B} $n \times n$ -es mátrixok, \mathbf{c} n dimenziós vektor, az ún. forrásvektor. \mathbf{A} , \mathbf{B} elemei és \mathbf{c} a \mathbf{v} komponenseinek függvényei. Ha a (2.8) egyenletrendszernek vannak komplex karakterisztikái, akkor a megoldás nem stabil, a kezdeti értékekben elkövetett kis hiba a megoldás során korlátlanul nőhet.

Jelöljük a (2.8) egyenletrendszer karakterisztikus gyökeit μ -vel, ezeket a

$$(2.9) \quad \det(\mathbf{A} - \mu \mathbf{B}) = 0$$

egyenletrendszer határozza meg.

Vizsgáljuk (2.8) lokális lineáris stabilitását, azaz vizsgáljuk, hogy ha a megoldást egy időpontban kis $\delta \mathbf{v}$ értékkel perturbáljuk, akkor ez a perturbáció hogyan terjed a megoldás során.

Ehhez (2.8)-ban írjunk \mathbf{v} helyébe $(\mathbf{v} + \delta \mathbf{v})$ -t, s a kapott eredményt linearizáljuk $\delta \mathbf{v}$ -re vonatkozóan. Így az alábbi egyenletrendszerhez jutunk, amely a \mathbf{v} , $\frac{\partial \mathbf{v}}{\partial t}$ és $\frac{\partial \mathbf{v}}{\partial x}$ pillanatnyi értékéhez tartozó perturbáció nélküli megoldás körüli kis $\delta \mathbf{v}$ perturbációt írja le:

$$(2.10) \quad \mathbf{A}(\mathbf{v}) \frac{\partial}{\partial t} (\delta \mathbf{v}) + \mathbf{B} \frac{\partial}{\partial x} (\delta \mathbf{v}) + \left(\delta \mathbf{v} \frac{\partial \mathbf{A}}{\partial \mathbf{v}} \right) \frac{\partial \mathbf{v}}{\partial t} + \left(\delta \mathbf{v} \frac{\partial \mathbf{B}}{\partial \mathbf{v}} \right) \frac{\partial \mathbf{v}}{\partial x} + \delta \mathbf{v} \frac{\partial \mathbf{c}}{\partial \mathbf{v}} = 0$$

Ezeket a pillanatnyi értékeket konstansoknak tekintve a $\delta \mathbf{v}$ perturbációra a következő alakú megoldást keressük:

$$(2.11) \quad \delta \mathbf{v} = \delta \mathbf{v}_0 \exp [i(kx - \omega t)].$$

Ezt (2.10)-be írva kapjuk, hogy

$$(2.12) \quad -i\omega \mathbf{A}(\mathbf{v}) \delta \mathbf{v}_0 + ik \mathbf{B}(\mathbf{v}) \delta \mathbf{v}_0 + \left(\delta \mathbf{v}_0 \frac{\partial \mathbf{A}}{\partial \mathbf{v}} \right) \cdot \frac{\partial \mathbf{v}}{\partial t} + \left(\delta \mathbf{v}_0 \frac{\partial \mathbf{B}}{\partial \mathbf{v}} \right) \frac{\partial \mathbf{v}}{\partial x} + \delta \mathbf{v}_0 \cdot \frac{\partial \mathbf{c}}{\partial \mathbf{v}} = 0.$$

Ennek a $\delta \mathbf{v}_0$ -ban homogén lineáris egyenletrendszernek akkor van nemtriviális megoldása, ha

$$(2.13) \quad \det(-i\omega \mathbf{A} + ik \mathbf{B} + \mathbf{D}) = 0,$$

ahol

$$(2.14) \quad \mathbf{D} = \left(\frac{\partial \mathbf{A}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial t} \right)^T + \left(\frac{\partial \mathbf{B}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial x} \right)^T + \left(\frac{\partial \mathbf{c}}{\partial \mathbf{v}} \right)^T.$$

Ha $\omega = \omega(k) \neq 0$, akkor (2.13) a következő alakba írható:

$$(2.15) \quad \det \left(\mathbf{A} - \frac{k}{\omega} \mathbf{B} + \frac{i}{\omega} \mathbf{D} \right) = 0.$$

Adott valós k értékhez különböző $\omega(k)$ komplex gyökök tartozhatnak. A (2.11) összefüggés alapján a stabilitásnak az a feltétele, hogy minden $\omega(k)$ gyökre $\text{Im}(\omega) \leq 0$ legyen. Ekkor lesz a perturbáció időbeli növekedése korlátos.

Ha $D=0$, akkor az ω értékek (2.9) és (2.15) összevetéséből a $\mu = \frac{k}{\omega}$ összefüggésből meghatározhatók. Ha minden μ gyök valós, akkor ezzel minden ω is valós lesz, tehát teljesül, hogy minden ω -ra $\text{Im}(\omega) \leq 0$. Ha van komplex μ gyök, akkor az komplex konjugáltjával együtt szerepel, s ekkor az egyik $\omega(k)$ -ra $\text{Im}(\omega) > 0$ lesz, vagyis a perturbáció korlátlanul nő.

Ha $D \neq 0$, akkor vizsgáljuk a $k \rightarrow \infty$ és $\omega \rightarrow \infty$ esetet. Feltesszük, hogy $\frac{k}{|\omega(k)|}$ határértéke véges, ha $k \rightarrow \infty$. (A valóságban mindig ez áll fenn.) Ekkor *Osztrovszkij tétele*ből [11] következik, hogy

$$(2.16) \quad \lim_{k \rightarrow \infty} \left| \mu - \frac{k}{\omega} \right| = 0.$$

Ha μ komplex, azaz $\mu = \mu_R + i\mu_I$, ($\mu_I \neq 0$), akkor (2.16)-ból következik, hogy

$$(2.17) \quad \lim_{k \rightarrow \infty} \left[\frac{-\text{Im}(\omega)k}{|\omega|^2} \right] = \mu_I.$$

Mivel $\frac{k}{|\omega|}$ -nak véges a határértéke (még hozzá pozitív vagy nulla), ha $k \rightarrow \infty$, ezért $\frac{|\omega|^2}{k} \rightarrow +\infty$, ha $k \rightarrow \infty$. Mivel a komplex karakterisztikák konjugáltjakkal együtt fordulnak elő, ezért feltehetjük, hogy $\mu_I < 0$. Ekkor (2.17)-ből az következik, hogy $\text{Im}(\omega) \rightarrow +\infty$, ha $k \rightarrow \infty$. Ezzel kaptuk, hogy nagyon kis hullámhosszokra instabilitás lép fel, sőt a növekedés nem korlátos, ha $k \rightarrow \infty$.

A fenti tételeből kiderül, hogy a (2.1)–(2.4) egyenleteket tartalmazó kezdeti érték feladatban a komplex karakterisztikák miatt a megoldás nem függ folytonosan a kezdeti értékektől.

3. A numerikus megoldás

A folyadékaramlási feladatok megoldására általánosan elterjedt módszer a véges differenciák módszere. Ennek előnye, hogy univerzálisan alkalmazható, viszonylag egyszerűen kiterjeszthető több dimenziós feladatok megoldására, a kezdeti és peremfeltételek változása nem befolyásolja döntően az eljárást.

Alkalmazásukkor fellép a pontosság, stabilitás és konvergencia kérdése [9].

Az előző pontból kiderült, hogy az alapvető két-közeg modellekben a megoldás nem függ folytonosan a kezdeti értékektől, a kiinduló parciális differenciálegyenlet-rendszer nem stabil. Ennek a következménye, hogy egyetlen vele konzisztens véges differencia séma sem lehet stabil [9].

A következőkben azt vizsgáljuk, hogy hogyan lehet mégis a gyakorlat számára megfelelő eredményeket kapni. A következő elrendezés alkalmazásával gyakran talál-

kozunk az irodalomban [3, 4, 7]

$$(3.1) \quad \frac{\alpha_j^{n+1} \varrho_{v,j}^{n+1} - \alpha_j^n \varrho_{v,j}^n}{\Delta t} + \frac{\alpha_j^n \varrho_{v,j}^n v_{v,j+1/2}^{n+1} - \alpha_{j-1}^n \varrho_{v,j-1}^n v_{v,j-1/2}^{n+1}}{\Delta x} = 0,$$

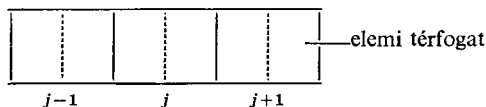
$$(3.2) \quad \frac{(1 - \alpha_j^{n+1}) \varrho_{l,j}^{n+1} - (1 - \alpha_j^n) \varrho_{l,j}^n}{\Delta t} + \frac{(1 - \alpha_j^n) \varrho_{l,j}^n v_{l,j+1/2}^{n+1} - (1 - \alpha_{j-1}^n) \varrho_{l,j-1}^n v_{l,j-1/2}^{n+1}}{\Delta x} = 0,$$

$$(3.3) \quad \alpha_j^n \varrho_{v,j}^n \left[\frac{v_{v,j+1/2}^{n+1} - v_{v,j+1/2}^n}{\Delta t} + v_{v,j+1/2}^n \frac{v_{v,j+1/2}^n - v_{v,j-1/2}^n}{\Delta x} \right] + \alpha_j^n \frac{P_{j+1}^{n+1} - P_j^{n+1}}{\Delta x} = K_{j+1/2} [v_{l,j+1/2}^{n+1} - v_{v,j+1/2}^{n+1}],$$

$$(3.4) \quad (1 - \alpha_j^n) \varrho_{l,j}^n \left[\frac{v_{l,j+1/2}^{n+1} - v_{l,j+1/2}^n}{\Delta t} + v_{l,j+1/2}^n \frac{v_{l,j+1/2}^n - v_{l,j-1/2}^n}{\Delta x} \right] + (1 - \alpha_j^n) \frac{P_{j+1}^{n+1} - P_j^{n+1}}{\Delta x} = K_{j+1/2} [v_{v,j+1/2}^{n+1} - v_{l,j+1/2}^{n+1}].$$

A (3.1)–(3.4) séma ilyen felírását stabilitási megfontolások és gyakorlati számításai tapasztalatok indokolják.

Az egydimenziós modellben a séma felírásához a reaktor zónáját (a valóságos értelmezési tartományt) elemi térfogatokra osztjuk az x -tengely mentén, s a változók értékét egy-egy ilyen elemi térfogaton vett átlag adja. Az elemi térfogatok elrendezését az alábbi ábra szemlélteti:



A térfogatoskat a folytonos vonalak jelentik, a szaggatott vonal a középpontjukon át húzódik. Az elemi térfogatok középpontjai a térbeli rácspontok. A tömeg és energia egyenleteket a rácspontok felezőpontja fölött differenciáljuk, az impulzus egyenleteket pedig a rácspontok fölött.

A (3.1)–(3.4) egyenletekben az n és $n+1$ indexek jelölik a változók értékeit az egymást követő időpontokban, Δt az időlépés, Δx a rácspontok távolsága. Az egyenletek szemi-implicit rendszert alkotnak: egyes változók a régi, mások az új időpontban vett értékükkel szerepelnek.

A tömegmegmaradási egyenletekben a sebességet tartalmazó tagok, az impulzus egyenletekben a nyomásgradiens és az impulzus-átadást tartalmazó tagok tartalmaznak olyan változót, melynek értékét az új időpontban vesszük. Az explicit és implicit kifejezések megválasztását a következő okok indokolják:

1) Lassú és gyors tranziensek leírására egyaránt szükség lehet. Gyors tranziensek esetén kis időlépésekre van szükség, hogy az áramlás alakulását minél pontosabban leírassuk. Másrészt viszont lassú folyamatoknál csökkenthetjük a számítási igényt, ha lehetőség van nagy időlépések vételére. Teljesen explicit séma esetén az időlépés

felső korlátját a *Courant-féle* $\left| (v \pm c) \frac{\Delta t}{\Delta x} \right| < 1$ stabilitási feltétel adja, ahol c a hang terjedési sebességét jelenti. Ha azonban azokat a tagokat, melyek a hang terjedését magukba foglalják implicit módon differenciáljuk, akkor ez a korlát

$$\max \left(\left| v_v \frac{\Delta t}{\Delta x} \right|, \left| v_l \frac{\Delta t}{\Delta x} \right| \right) < 1$$

-re növelhető. (Ilyenek a (3.3) és (3.4) egyenletben a nyomásgradienst tartalmazó tagok.)

2) Az időlépés korlátját növelhetnénk a további tagok implicit kezelésével. Az így kapott teljesen implicit séma megoldásához viszont megnőne a szükséges iterációk száma, jóval komplikáltabbá válna a kapott algebrai egyenletrendszer megoldása, jelesen megnőne a gépidő igény.

Az α , q_v , q_l , P változók értékét a rácspontokban, v_v és v_l értékét pedig csak felezőpontjaiban definiáljuk. A séma felírásakor a változók egy része olyan pontokban is szerepel, ahol értéke nincs definiálva. Ahhoz, hogy az egyenletek minden rácspontban értelmezve legyenek, meg kell adni olyan összefüggéseket, melyek a változóknak az egyes elemi térfogatok középpontjában és a határan felvett értékeit kapcsolják össze. Erre a célra szolgál az alábbi összefüggés:

3.1. DEFINÍCIÓ. Az f változó értékét a fél rácspontokban a következő, donor összefüggés adja meg:

$$(3.5) \quad f_{j+1/2} = \begin{cases} f_{j+1}, & \text{ha } v_{j+1/2} < 0 \\ f_j, & \text{ha } v_{j+1/2} \geq 0 \end{cases}$$

Az elnevezést az indokolja, hogy a változó értékét egy elemi térfogat falán abban az elemi térfogatban felvett értéke határozza meg, amely felől a folyadék folyik, vagyis az előző térfogat donorként viselkedik.

A (3.1)–(3.4) egyenletek felírásában már szerepelnek a donor összefüggések is azzal a feltételezéssel, hogy $v_v, v_l > 0$. A tömegmegmaradási egyenletekben a hely szerinti deriváltat tartalmazó tagokban, az impulzus egyenletekben pedig a konvektív tagokban használjuk a donor összefüggéseket. Az így megadott donor összefüggések stabilizáló hatással vannak a véges differencia egyenletrendszer megoldására [2].

4. A numerikus séma vizsgálata

A séma stabilitását *Neumann módszerével* vizsgáljuk. Az eljárást NEUMANN eredetileg lineáris konstans együtthatós egyenletekre dolgozta ki [1].

Lényege, hogy a véges differencia egyenletben az ismeretlen függvény j -edik rácspontban és n -edik időpontban vett értéke helyébe $\lambda^n e^{imj\Delta x}$ -et helyettesítünk, ahol $m = \frac{\pi}{k\Delta x}$, $k = 1, 2, \dots, N$ és N a rácshintervallumok száma. Vizsgáljuk, hogy ez λ milyen értéke mellett elégíti ki az egyenletet.

A stabilitás feltétele, hogy $\max_{(m)} |\lambda(m)| \leq 1$ legyen.

A gyakorlatban ezzel a módszerrel hasznos eredmények nyerhetők nem konstans együtthatós és nem lineáris esetekben is. Ilyenkor a megoldást egy pont körül lineárizáljuk, s a vizsgált tartományt olyan kis részekre bontjuk, ahol az együtthatók elhanyagolható hibával konstansnak tekinthetők. Ezzel a véges differencia egyenlet lokális lineáris stabilitása megvizsgálható, amiből az eredeti rendszerre vonatkozóan lehet következtetéseket levonni.

Ha az eljárást egyenletrendszerre alkalmazzuk, akkor a változók időbeli növekedését az amplifikációs mátrix írja le [5], ennek segítségével a *Neumann-féle stabilitás feltétel* azt mondja ki, hogy az amplifikációs mátrix sajátértékeinek abszolút értéke ne legyen nagyobb 1-nél [6].

Vizsgáljuk meg a (3.1)–(3.4) véges differencia egyenletrendszer lokális lineáris stabilitását [10] alapján. Ehhez először kifejezzük a bennük előforduló differenciákat a q_v , α , v_v , v_l négy alapváltozó differenciái segítségével. Ezután ezek együtthatóit konstansnak tekintve vizsgáljuk egy *Fourier komponens* növekedését.

Tekintsük először a $K=0$ esetet, tehát amikor nincs impulzus-átadás a fázisok közt. Ekkor a következő amplifikációs mátrixot kapjuk:

$$(4.1) \quad \begin{bmatrix} \alpha(\lambda - 1 + \tilde{v}_v) & q_v(\lambda - 1 + \tilde{v}_v) & \lambda \frac{\Delta t}{\Delta x} \alpha q_v im' & 0 \\ 0 & -q_l(\lambda - 1 + \tilde{v}_l) & 0 & \lambda \frac{\Delta t}{\Delta x} (1 - \alpha) q_l im' \\ \lambda \frac{\Delta t}{\Delta x} c_v^2 im' & 0 & q_v(\lambda - 1 + \tilde{v}_v) & 0 \\ \lambda \frac{\Delta t}{\Delta x} c_v^2 im' & 0 & 0 & q_l(\lambda - 1 + \tilde{v}_l) \end{bmatrix},$$

ahol az oszlopok balról jobbra haladva a q_v , α , v_v , v_l változóknak felelnek meg; m a hullámszám $m = \frac{\pi}{k \Delta x}$ és $m' = 2 \sin \frac{1}{2} m \Delta x$,

$$\tilde{v}_v = v_v \frac{\Delta t}{\Delta x} [1 - \exp(-im \Delta x)] \quad \text{és} \quad \tilde{v}_l = \frac{v_l}{v_v} \tilde{v}_v.$$

A (4.1) mátrix sajátértékei a

$$(4.2) \quad \lambda^2 c_k^2 \left[(\lambda - 1 + \tilde{v}_l)^2 + \frac{(1 - \alpha) q_v}{\alpha q_l} (\lambda - 1 + \tilde{v}_v)^2 \right] + (\lambda - 1 + \tilde{v}_v)^2 (\lambda - 1 + \tilde{v}_l)^2 = 0$$

egyenletnek tesznek eleget, ahol

$$c_k = c_v \frac{\Delta t}{\Delta x} 2 \sin \frac{\pi}{2k}.$$

Vizsgáljuk először a rendszer viselkedését nagy frekvenciákra (kis k -ra). A fenti differencia séma használata elsősorban akkor előnyös, ha Δt megközelíti felső korlát-

ját $\left(\text{azaz } \Delta t \approx \min \left(\frac{\Delta x}{v_v}, \frac{\Delta x}{v_l} \right) \right)$, és lényegesen nagyobb mint $\frac{\Delta x}{c_v}$, vagyis $v_v, v_l \ll c_v$.

Ekkor $c_v \frac{\Delta t}{\Delta x} \gg 1$ és így kis k -ra $c_k \gg 1$.

Ekkor (4.2) két gyöke közel $c_k^{-1} \ll 1$ abszolút értékű, a másik két gyöke pedig közelítőleg kielégíti a

$$(4.3) \quad \lambda - 1 + \tilde{v}_l \cong \pm i\varepsilon(\lambda - 1 + \tilde{v}_v)$$

összefüggést, ahol

$$\varepsilon^2 = \frac{(1 - \alpha)\varrho_v}{\alpha\varrho_l}.$$

Ebből adódik, hogy

$$(4.4) \quad \lambda \cong 1 - \tilde{v}_l \left(1 \pm i\varepsilon \frac{v_v}{v_l} \right) (1 \pm i\varepsilon)^{-1}.$$

Ezek a pontok a $v_l \frac{\Delta t}{\Delta x}$ sugarú és az 1 pontot érintő körön helyezkednek el a komplex síkon, amely az 1 pont körül el van forgatva $\pm \arctan \left(\varepsilon \frac{v_v}{v_l} \right)$ szöggel, majd visszafelé $\pm \arctan \varepsilon$ szöggel. A nagyítás (nyújtás) korlátja $1 + \left| \frac{v_v}{v_l} \right|$. Ezen a körön levő pontok kis k -ra (tehát az 1-től legtávolabb eső pontokra) nem esnek az egységkörön kívülre, ha $v_l \frac{\Delta t}{\Delta x} < 1$ és $v_v \frac{\Delta t}{\Delta x} < 1$.

Ha tehát az időlépés nem lépi túl az ebből eredő $\Delta t < \min \left(\frac{\Delta x}{v_l}, \frac{\Delta x}{v_v} \right)$ korlátot, akkor a legnagyobb frekvenciájú komponensek (azaz amelyek hullámhossza a Δx rácstávolság kisszámú többszöröse) nem nőnek exponenciálisan. (Még az impulzusátadás nélküli két-közeg modellben sem.)

Az alacsony frekvenciájú komponensek vizsgálatához tekintsük (4.2)-t nagy k -ra. Mivel $k \rightarrow \infty$ esetén $c_k \rightarrow 0$, ezért a (4.2) gyökei $\lambda = 1 - \tilde{v}_v$ -hez és $\lambda = 1 - \tilde{v}_l$ -hez tartanak (mindkettőhöz kétszeresen). Pontosabb vizsgálatához legyen $k \rightarrow \infty$ -re

$$\lambda = 1 - \tilde{v}_l + \delta.$$

Ezt beírva (4.2)-be δ -ra kapunk egy összefüggést, melyből elhagyjuk a δ -ban másodfokúnál magasabb fokú tagokat. Az eredményül kapott δ -ban másodfokú polinom gyökei a következő összefüggést elégítik ki

$$(4.5) \quad \frac{\delta}{(1 - \tilde{v}_l)(\tilde{v}_v - \tilde{v}_l)} \cong \frac{\varepsilon^2 \pm i\varepsilon \left(1 - \frac{(v_v - v_l)^2}{c_v^2} \right)^{1/2}}{1 + \varepsilon^2 - \frac{(v_v - v_l)^2}{c_v^2}}.$$

Mivel a (3.1)–(3.4) séma használatakor feltehetjük, hogy $|v_v - v_l| \ll c_v$, így

$$(4.6) \quad \lambda \approx (1 - \tilde{v}_l) \left[1 - i(\tilde{v}_v - \tilde{v}_l) \frac{\pi \Delta t}{k \Delta x} \varepsilon \left(\frac{\varepsilon \pm i}{1 + \varepsilon^2} \right) \right].$$

Mivel $|1 - \tilde{v}_l| = 1 - O(k^{-2})$ és ezt még egy olyan mennyiséggel szorozzuk, melynek nagysága $1 + O(k^{-1})$, ahol $O(k^{-1})$ pozitív lesz a (4.6)-ban szereplő \pm egyike esetén, ezért az egyik λ gyök szükségképpen az egységkörön kívülre esik nagy k -ra.

Hasonló eredményhez jutunk, ha a $\lambda = 1 - \tilde{v}_v + \delta$ kifejezést vizsgáljuk. Tehát a (3.1)–(3.4) séma $K=0$ esetén a nagy k -hoz tartozó hullámhosszakra instabil lesz. k értéke úgy növelhető, ha növeljük az osztáspontok N számát. Ez kétféleképpen történhet. Először úgy, hogy Δx -et rögzítve a fizikai hosszat növeljük. A másik lehetőség, hogy a fizikai hosszúság változatlanul hagyása mellett csökkentjük Δx -et és Δt -t, azaz finomítjuk a rácsot. Erre $K=0$ esetén azért van lehetőség, mert ekkor az egyenletek invariánsak az idő és térbeli egyenletes kicsinyítéssel szemben. $K=0$ -ra tehát azt kaptuk, hogy a (3.1)–(3.4) egyenletekkel végzett számítások kevés térbeli osztáspont esetén „jól viselkedő” megoldást adnak. A rács finomításakor azonban instabilitások lépnek fel, mégpedig a Δx nagyszámú többszöröseinek megfelelő hullámhosszakra.

Vizsgáljuk meg ezután a (3.1)–(3.4) rendszert $K>0$ esetén. Az impulzus egyenletekben a sebességek implicit kifejezésekként szerepelnek, hogy az impulzusok közti szoros kapcsolat miatt ne legyen szükség Δt további korlátozására.

A lineáris stabilitásvizsgálathoz felírt amplifikációs mátrix (4.1)-től a jobb alsó sarokban tér el, melynek alakja most a következő lesz:

$$(4.7) \quad \begin{bmatrix} \varrho_v[\lambda(1-\kappa) - 1 + \tilde{v}_v] & -\lambda\varrho_v\kappa \\ -\lambda\varrho_v\kappa & \varrho_l(\lambda - 1 + \tilde{v}_l) + \lambda\varrho_v\kappa \end{bmatrix},$$

ahol $\kappa = K \frac{\Delta t}{\varrho_v}$ (az impulzus-átadásban akkor érdemes az implicit sebességeket alkalmazni, ha $\kappa > 1$).

A most kapott karakterisztikus polinomra ismét meg kell vizsgálni, hogy teljesül-e rá a stabilitás feltétele, azaz minden λ -ra $|\lambda| \leq 1$.

Vizsgáljuk először a nagy frekvenciákat, azaz legyen k kicsi. Feltéve, hogy $\kappa \gg 1$ és az előző gondolat menetéhez hasonlóan $c_k \gg 1$ azt kapjuk, hogy a karakterisztikus polinom két gyökének nagysága közel c_k^{-1} , egyé nagyjából κ^{-1} és a negyedik közelítőleg egyenlő

$$1 - \frac{\tilde{v}_l + \varepsilon^2 \tilde{v}_v}{1 + \varepsilon^2} \text{-tel, ahol } \varepsilon^2 = \frac{(1 - \alpha)\varrho_v}{\alpha\varrho_l}.$$

Ez is az egységkörön belülre esik, ha $\Delta t \leq \min \left(\frac{\Delta x}{v_v}, \frac{\Delta x}{v_l} \right)$ teljesül.

Igaz az is, hogy kis k -ra és nagyon kis κ -ra mind a négy gyök az egységkörön belül van, mert $\kappa=0$ -ra ezt az előbbiekben beláttuk ($K=0$ eset) és a gyökök folytonosan függenek κ -tól. Ezzel nagy κ -ra (amikor leginkább szükség van az impulzus átadás implicit kezelésére) és nagyon kis κ -ra bizonyítottuk az állítást, de igaznak tehetjük fel a közbülső κ -ra is, mivel az implicit módon kezelt csillapítást tartalmazó egyenletekben nyilván még annyira se növekednek a Fourier komponensek, mint az ilyeneket nem tartalmazókban.

Ezzel beláttuk, hogy nagy frekvenciára a komponensek stabilak lesznek a (3.1)–(3.4) egyenletrendszer megoldásában, ha $K>0$.

Az alacsony frekvenciák vizsgálatához ismét a $k \rightarrow \infty$ esetet kell tekinteni. Az $N \rightarrow \infty$ kétféle értelmezése most különböző megoldáshoz vezet a $K(v_v - v_l)$ nem differenciális tag jelenléte miatt.

Ha N növelése a rács finomítását jelenti, akkor $\kappa = K \frac{\Delta t}{\varrho_v}$ -nek nullához kell tartania. A $K(v_v - v_l)$ tag implicit módon fizikai hosszat tartalmaz. Legyen $k \rightarrow \infty$ úgy, hogy közben κ értékét megtartjuk. Ez úgy értelmezhető, hogy a rácsot finomítjuk és feltesszük, hogy egy bizonyos ponton $(\Delta x)^{-1}$ és $(\Delta t)^{-1}$ minden magasabb rendű kifejezése elhanyagolható κ kivételével, vagy pedig Δx -et rögzítjük és a fizikai hosszát növeljük.

Fix κ -ra és $k \rightarrow \infty$ -re határértékként a következő egyenletet kapjuk:

$$(4.8) \quad (\lambda - 1 + \tilde{v}_v)(\lambda - 1 + \tilde{v}_l)[(\kappa\lambda + \lambda - 1)(\varrho_l(\lambda - 1) + \lambda^2\varrho_v\kappa) + \lambda^2\varrho_v\kappa^2] = 0.$$

A szögletes zárójelben levő másodfokú polinom gyökeiről megmutatható, hogy valóságosak, pozitívak és egynél szigorúan kisebbek. Ezért elég nagy k -ra a megfelelő gyököknek az egységkörön belül kell lenniük.

A másik két gyök vizsgálatához az előbbiekhöz hasonlóan legyen $\lambda = 1 - \tilde{v}_v + \delta$, ezt helyettesítsük be az eredeti polinom egyenletbe, s csak a δ -ban elsőfokú tagokat hagyjuk meg. Hanyagoljuk el továbbá k^{-1} magasabb fokú kifejezéseit. Végül feltételezve, hogy

$$(4.9) \quad \varrho_l |\tilde{v}_l - \tilde{v}_v| \ll 2\varrho_v\kappa$$

a maradék gyök közelítőleg

$$\lambda = (1 - \tilde{v}_v) \frac{\kappa + c_k^2}{\kappa + 2c_k^2},$$

ami pedig az egységkörön belül van. Ugyanezt az eredményt kapjuk $\lambda = 1 - \tilde{v}_l + \delta$ helyettesítéssel.

Ezzel beláttuk, hogy a hosszú hullámhosszú komponensek is stabilak. A (4.9) feltétel teljesülése nagyon fontos ahhoz, hogy nagy hullámhosszakra is jó viselkedést kapjunk. Jelentése az, hogy az impulzus-átadás mértékének egy bizonyos értéknél nagyobboknak kell lennie ahhoz, hogy ne lépjenek fel korlátlanul növvő hosszú hullámhosszú komponensek. Ez a minimális érték a rácsméret és az impulzus-átadás fizikailag értelmezhető kapcsolatából adódik. Nagy k -ra ugyanis (4.9) az alábbi összefüggéssel ekvivalens:

$$(4.10) \quad \frac{K(k\Delta x)}{\varrho_v} \gg |v_v - v_l| \frac{\varrho_l}{\varrho_v}.$$

K viszont implicit módon tartalmaz egy fizikai hosszat, pl. a [10]-ben javasolt alakja buboréknak folyadékban való áramlása esetén:

$$(4.11) \quad K \cong C_D \varrho_v \frac{|v_v - v_l|}{r},$$

ahol r a buboréksugár, C_D pedig a közegellenállási tényező. Ennek felhasználásával (4.9) a következő alakba írható:

$$(4.12) \quad k\Delta x \gg \frac{\varrho_l}{\varrho_v C_D} r,$$

azaz a $k\Delta x$ hullámhosszú komponens nem fog korlátlanul nőni, ha ez a hullámhossz nagyobb, mint az egyes buborékok sugarának bizonyos számú többszöröse.

Ezzel beláttuk, hogy a (3.1)–(3.4) differencia egyenletekkel végzett számításokban a nagy frekvenciájú komponensek nem növekvők és azok a nagy k -hoz tartozó kis frekvenciájú komponensek, melyek egy adott rácsra kielégítik (4.10)-et, szintén jól viselkednek. Ez azonban még nem biztosítja, hogy egyáltalán ne legyen növekvő komponens. Ehhez (4.10) teljesítésülését kellene megkövetelni egészen $k=1$ -től látszólag, ekkor áthidalnánk a kis és nagy frekvenciák közti rést. De (4.10) csak nagy k -kra érvényes.

A vizsgálatból kiderült, hogy ha $K\Delta x$ elég nagy, akkor az impulzus-átadás stabilizálja azokat a nagy hullámhosszakat, melyekhez egyébként korlátlanul növekvő komponensek tartoznának.

A gyakorlatban a (3.1)–(3.4) sémához hasonló sémákkal végzett számítások azt mutatják, hogy minden hullámhosszra jól viselkedő (azaz nem korlátlanul növekvő) megoldást kapunk. Ha viszont csökkentjük K értékét, akkor hosszú hullámhosszokra növekvő amplitúdójú oszcilláció léphet fel [10].

Összegezve tehát megmutattuk, hogy a kétfázisú áramlás (2.1)–(2.4) két-közeg modelljével végzett véges differencia számítások jól viselkedő megoldást adnak, ha a fázisok közti impulzus-átadás elég nagy és a térbeli rács nem túl finom. Ez annak ellenére igaz, hogy a két-folyadék differenciálegyenlet-rendszernek vannak komplex karakterisztikái, ezért a megoldás nem függ folytonosan a kezdeti értékektől, és az elegendően finom rácson végzett számítások emiatt szükségképpen instabilak. Származtattunk egy fizikailag értelmezhető kritériumot a rács méretére vonatkozólag. Ennek jelentése, hogy a közelítő számítások mindaddig jók lehetnek, amíg nem próbáljuk olyan jelenségekre alkalmazni, melyek finomabbak az impulzus-átadási kifejezésben implicit módon tartalmazott méretnél.

Másrészt viszont, ha $K=0$ (azaz nincs impulzusátadás, tehát szeparált az áramlás), akkor a térfogattörtek változása a gőzfolyadék érintkezési felület változásának eredménye. Ekkor minden olyan közelítő megoldás, amelyben megengedett a térfogattörtek változása, szükségképpen finom megoldás. Ebben az esetben a felületi feszültség figyelembe vételével stabil feladatot kapunk. A következő pontban [8] alapján ezt a lehetőséget vizsgáljuk meg.

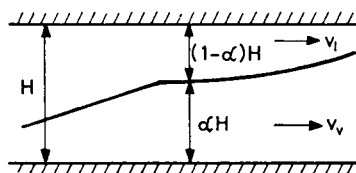
5. A modell módosítása finom megoldás esetére

$K=0$ esetében a (2.1)–(2.4) egyenletrendszer a párhuzamos síkok között történő szeparált kétfázisú áramlást írja le azonos nyomásmezőben. A továbbiakban ezt alapegyenletrendszernek nevezzük. A 2) pontban megmutattuk, hogy ennek két komplex karakterisztikus gyöke van. Emiatt a megoldások nagyon rövid hullámhosszokra instabilak, nem függnek folytonosan a kezdeti feltételektől. A komplex karakterisztikájú egyenletek helyesen írhatják le a hosszú hullámhosszú fizikai instabilitásokat, a hullámhossz csökkentésével azonban a hiba egyre növekszik, s elég rövid hullámhosszokra nem-fizikai jellegű instabilitások lépnek fel. A kutatók egy jelentős része szerint a fizikai jelenségeket le kell tudni írni korrekt feladatokkal. Ha egy feladat nem ilyen, akkor ennek oka, hogy olyan jelenséget hanyagoltunk el, amely valamilyen szempontból lényeges. Ha ezeket beépítjük a feladatba, akkor valósakká válnak a ka-

rakterisztikák. A komplex karakterisztikák okozta nehézségek kiküszöbölhetők, ha olyan ismert fizikai hatásokat foglalunk a modellbe, melyek rövid hullámhosszaknál jelentőssé válnak, ugyanakkor hosszú hullámhosszaknál nem változtatják meg a rendszer viselkedését.

Ilyen rövid hullámhosszú fizikai hatás a felületi feszültség, vagy a viszkozitás. Megmutatjuk, hogy ennek beépítésével a karakterisztikák valósakká válnak, a rövid hullámhosszú perturbációk stabilizálódnak. Ugyanakkor a fizikai természetű (*Helmholtz-féle*) hosszú hullámhosszú instabilitások megmaradnak és számolhatók a modellel.

Az áramlás szerkezetét az alábbi ábra mutatja:



1. ábra

A (2.1)–(2.4) egyenletrendszerbe úgy tudjuk beépíteni a felületi feszültséget, ha feltételezzük, hogy a fázisok kissé különböző nyomáson vannak, ezeket a nyomásokat a felületi feszültség kifejezése kapcsolja össze. Az alap-egyenletrendszerből így a következő egyenletrendszerhez jutunk

$$(5.1) \quad \frac{\partial}{\partial t}(\alpha \varrho_v) + \frac{\partial}{\partial x}(\alpha \varrho_v v_v) = 0$$

$$(5.2) \quad \frac{\partial}{\partial t}[(1-\alpha) \varrho_l] + \frac{\partial}{\partial x}[(1-\alpha) \varrho_l v_l] = 0$$

$$(5.3) \quad \alpha \varrho_v \frac{\partial v_v}{\partial t} + \alpha \varrho_v v_v \frac{\partial v_v}{\partial x} + \alpha \frac{\partial P_1}{\partial x} = 0$$

$$(5.4) \quad (1-\alpha) \varrho_l \frac{\partial v_l}{\partial t} + (1-\alpha) \varrho_l v_l \frac{\partial v_l}{\partial x} + (1-\alpha) \frac{\partial P_2}{\partial x} = 0$$

$$(5.5) \quad P_1 - P_2 = -\sigma H \frac{\partial^2 \alpha}{\partial x^2},$$

ahol σ a felületi feszültség, H a síkok egymástól való távolsága. (5.5) a $P_1 - P_2 = \frac{\sigma}{R}$ összefüggésből származik, ahol R az érintkezési felület görbületi sugara, melyet az $\frac{1}{R} = -H \frac{\partial^2 \alpha}{\partial x^2} \left[1 + H^2 \left(\frac{\partial \alpha}{\partial x} \right)^2 \right]^{-3/2}$ összefüggés ad meg. Ebben feltesszük, hogy $H^2 \left(\frac{\partial \alpha}{\partial x} \right)^2 \ll 1$, különben a (5.1)–(5.4) egyenletek sem lennének érvényesek az egydimenziós változó keresztmetszetű áramlásra [8].

A továbbiakban a (5.1)–(5.5) rendszerre mint σ -egyenletrendszerre hivatkozunk.

ÁLLÍTÁS. A σ -egyenletrendszer karakterisztikus gyökei valósak.

A karakterisztikus gyökök meghatározásához a σ -egyenletrendszert először elsőrendűvé alakítjuk $\beta = \frac{\partial \alpha}{\partial x}$ változó bevezetésével. Ezzel az (5.5) egyenlet helyett az alábbi két egyenlet lép fel:

$$(5.6) \quad P_1 - P_2 + \sigma H \frac{\partial \beta}{\partial x} = 0,$$

$$(5.7) \quad \frac{\partial \alpha}{\partial x} - \beta = 0.$$

Feltesszük, hogy a sűrűségek csak a nyomás függvényei, azaz $\varrho_v = f_1(P_1)$ és $\varrho_l = f_2(P_2)$, ahol az f_1 és f_2 függvény ismert. Ezek segítségével ϱ_v és ϱ_l eliminálható a rendszerből.

Ezzel az egyenletrendszert felírhatjuk az alábbi mátrix formában:

$$(5.8) \quad \mathbf{A} \frac{\partial \mathbf{v}}{\partial t} + \mathbf{B} \frac{\partial \mathbf{v}}{\partial x} + \mathbf{c} = 0,$$

ahol \mathbf{v} a függő változók oszlopvektora

$$(5.9) \quad \mathbf{v} = (P_1, P_2, v_v, v_l, \alpha, \beta)^T.$$

Az \mathbf{A} és \mathbf{B} mátrix felépítése a következő:

$$\mathbf{A} = \begin{pmatrix} \frac{\alpha}{c_v^2} & 0 & 0 & 0 & \varrho_v & 0 \\ 0 & \frac{1-\alpha}{c_l^2} & 0 & 0 & -\varrho_l & 0 \\ 0 & 0 & \alpha \varrho_v & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\alpha) \varrho_l & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} \frac{\alpha v_v}{c_v^2} & 0 & \alpha \varrho_v & 0 & \varrho_v v_v & 0 \\ 0 & \frac{(1-\alpha) v_l}{c_l^2} & 0 & (1-\alpha) \varrho_l & -\varrho_l v_l & 0 \\ \alpha & 0 & \alpha \varrho_v v_v & 0 & 0 & 0 \\ 0 & 1-\alpha & 0 & (1-\alpha) \varrho_l v_l & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma H \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

ahol $c_v = \left(\frac{dP_1}{d\varrho_v} \right)^{\frac{1}{2}}$ és $c_l = \left(\frac{dP_2}{d\varrho_l} \right)^{\frac{1}{2}}$ a hang terjedési sebessége a megfelelő fázisban.

A karakterisztikus gyököket a

$$(5.10) \quad \det(A - \mu B) = 0$$

egyenlet adja, ezek a következők: kétszer 0, $(v_l \pm c_l)^{-1}$ és $(v_v \pm c_v)^{-1}$.

Ezzel beláttuk, hogy a felületi feszültség hozzávételével az alapegyenletrendszer karakterisztikái valósakká válnak.

Vizsgáljuk most meg a σ -egyenletrendszer lineáris stabilitását. Analitikusan ezt az inkompresszibilis áramlás esetére (ϱ_v és ϱ_l konstans) tudjuk elvégezni abban az esetben, ha a perturbáció nélküli megoldás stacionárius és izotróp. Ekkor a σ -egyenletrendszer alakja a következő lesz:

$$(5.11) \quad \frac{\partial \alpha}{\partial t} + v_v \frac{\partial \alpha}{\partial x} + \alpha \frac{\partial v_v}{\partial x} = 0$$

$$(5.12) \quad -\frac{\partial \alpha}{\partial t} - v_l \frac{\partial \alpha}{\partial x} + (1 - \alpha) \frac{\partial v_l}{\partial x} = 0$$

$$(5.13) \quad \frac{\partial v_v}{\partial t} + v_v \frac{\partial v_v}{\partial x} + \frac{\partial \tilde{P}_1}{\partial x} = 0$$

$$(5.14) \quad \frac{\partial v_l}{\partial t} + v_l \frac{\partial v_l}{\partial x} + \frac{\partial \tilde{P}_2}{\partial x} = 0$$

$$(5.15) \quad \varrho_v \tilde{P}_1 - \varrho_l \tilde{P}_2 = -\sigma H \frac{\partial^2 \alpha}{\partial x^2},$$

$$\text{ahol } \tilde{P}_1 = \frac{P_1}{\varrho_v} \text{ és } \tilde{P}_2 = \frac{P_2}{\varrho_l}.$$

A lineáris stabilitás vizsgálathoz minden u változó helyébe $u + \delta u \exp[i(kx - \omega t)]$ -t írunk, s a kapott egyenleteket linearizáljuk a δu perturbációkra vonatkozóan. Az izotróp és stacionárius perturbáció nélküli megoldás feltételezése azt jelenti, hogy u nem függ x -től és t -től, ezért a következő eredményt kapjuk:

$$(5.16) \quad (-i\omega + ikv_v)\delta\alpha + ik\alpha\delta v_v = 0$$

$$(5.17) \quad -(-i\omega + ikv_l)\delta\alpha + ik(1 - \alpha)\delta v_l = 0$$

$$(5.18) \quad (-i\omega + ikv_v)\delta v_v + ik\delta\tilde{P}_1 = 0$$

$$(5.19) \quad (-i\omega + ikv_l)\delta v_l + ik\delta\tilde{P}_2 = 0$$

$$(5.20) \quad -k^2\sigma H\delta\alpha + \varrho_v\delta\tilde{P}_1 - \varrho_l\delta\tilde{P}_2 = 0.$$

Az (5.16)–(5.20) egyenletrendszer homogén lineáris a $\delta\tilde{P}_1$, $\delta\tilde{P}_2$, δv_v , δv_l és $\delta\alpha$ változóiban. Ahhoz, hogy legyen nem-triviális megoldása az együttható mátrix determinánsának nullának kell lennie, azaz

$$(5.21) \quad \alpha\varrho_l(kv_l - \omega)^2 + (1 - \alpha)\varrho_v(kv_v - \omega)^2 - \alpha(1 - \alpha)\sigma Hk^4 = 0.$$

Adott k -ra (5.21)-ből az $\omega(k)$ gyökök meghatározhatók:

$$(5.22) \quad \omega = \frac{k\alpha(1-\alpha)}{(1-\alpha)\varrho_v + \alpha\varrho_l} \times \\ \times \left\{ \frac{\varrho_v v_v}{\alpha} + \frac{\varrho_l v_l}{1-\alpha} \pm \left[\frac{\sigma H k^2}{\alpha(1-\alpha)} [(1-\alpha)\varrho_v + \alpha\varrho_l] - \frac{\varrho_v \varrho_l}{\alpha(1-\alpha)} (v_v - v_l)^2 \right]^{\frac{1}{2}} \right\}.$$

A stabilitás feltétele, hogy minden $\omega(k)$ gyökre $\text{Im}(\omega) \leq 0$ legyen. Mivel ebben az esetben a komplex gyökök konjugáltjakkal együtt lépnek fel, így a stabilitás feltétele, hogy minden ω valós legyen, azaz

$$(5.23) \quad \sigma H k^2 [(1-\alpha)\varrho_v + \alpha\varrho_l] \geq \varrho_v \varrho_l (v_v - v_l)^2.$$

Ez az eredmény σ tetszőleges értékére, tehát $\sigma=0$ -ra is érvényes, a vizsgálat eredményei tehát mind az alap, mind a σ -egyenletrendszerre alkalmazhatók.

A (5.23) egyenlőtlenség $v_v = v_l$ esetén minden $\sigma \geq 0$ -ra fennáll. Ha tehát a fázisok azonos sebességgel mozognak, akkor az egyenletrendszer stabil lesz minden hullámhosszra.

A továbbiakban a $v_v \neq v_l$ esetet vizsgáljuk. Ha ekkor (5.23)-ba $\sigma=0$ -t írunk azt kapjuk, hogy az alap-egyenletrendszer minden hullámhosszra instabil. A σ -egyenletrendszerre ($\sigma > 0$) (5.23) azt fejezi ki, hogy a nagyon rövid hullámhosszak (nagy k) stabilak, a nagyon nagyok (kis k) viszont instabilak.

(5.23) segítségével megadhatunk egy k_c kritikus hullámszámot

$$(5.24) \quad k_c = |v_v - v_l| \left\{ \frac{\varrho_v \varrho_l}{\sigma H [(1-\alpha)\varrho_v + \alpha\varrho_l]} \right\}^{\frac{1}{2}}.$$

Ennek segítségével a $\frac{2\pi}{k_c}$ -nél hosszabb hullámhosszakra a σ -egyenletrendszer instabil, az ennél kisebb vagy vele egyenlő hullámhosszakra pedig stabil. (A hosszú hullámhosszakra kapott instabilitás a fizikailag értelmezett *Helmholtz-féle instabilitás*.)

Az (5.22) összefüggésből látható, hogy $k \ll k_c$ esetén az alap- és a σ -egyenletrendszer lényegében azonosan viselkedik, mivel a szögletes zárójelben levő első tag ekkor elhanyagolható a másodikhoz képest, mintha $\sigma=0$ lenne. Ezért a $\frac{2\pi}{k_c}$ -nél jóval nagyobb hullámhosszakhoz tartozó fizikai instabilitások mindkét rendszerrel helyesen számíthatók. Az ilyen jelenségek leírására tehát a komplex karakterisztikájú rendszer is alkalmazható.

A hullámhossz csökkenésével azonban az (5.22)-ben megadott ω egyre inkább más lesz $\sigma > 0$ és $\sigma=0$ esetén. $\frac{2\pi}{k_c}$ -nél kisebb hullámhosszakra pedig teljesen más eredményt ad a két egyenletrendszer; a σ -egyenletrendszer stabil viselkedést ír le, az alap-egyenletrendszer viszont instabil. A valósághoz természetesen a σ -egyenletrendszer áll közelebb, mert a felületi feszültség egy ténylegesen fellépő fizikai hatás, melyet egy pontosra törekvő fizikai modellezésnél figyelembe kell venni. Az alap egyenletrendszer $k > k_c$ -re pontatlan, rövid hullámhosszakra a valóságtól eltérő nem fizikai jellegű instabilitás lép fel a megoldásban.

Az (5.22) alapján felírható növekedési tényező:

$$(5.25) \quad G(k, t) \equiv |\exp[-i\omega(k)t]| = \exp[\operatorname{Im}(\omega)t].$$

Ha ω komplex, akkor (5.22)-ben a pozitív előjel felel meg a növekedésnek. $\sigma=0$ -t írva (5.22)-be az alap-egyenletrendszerre kapjuk, hogy

$$(5.26) \quad \operatorname{Im}(\omega) = \frac{k|v_v - v_l|}{(1-\alpha)\varrho_v + \alpha\varrho_l} [\alpha\varrho_v(1-\alpha)\varrho_l]^{\frac{1}{2}},$$

a σ -egyenletrendszerre pedig ($\sigma > 0$)

$$(5.27) \quad \operatorname{Im}(\omega) = \begin{cases} 0, & \text{ha } k \geq k_c \\ k \left[\frac{\alpha(1-\alpha)\sigma H}{(1-\alpha)\varrho_v + \alpha\varrho_l} (k_c^2 - k^2) \right]^{\frac{1}{2}}, & \text{ha } k < k_c. \end{cases}$$

[9] szerint a feladat akkor korrekt, ha egy adott t időpontban $G(k, t)$ a k hullámszám korlátos függvénye a $0 \leq k < \infty$ intervallumban, vagyis ha $\operatorname{Im}(\omega)$ korlátos k -ra vonatkozólag. Ez (5.27)-re igaz de (5.26)-ra nem.

Most vizsgáljuk meg, hogy milyen problémák lépnek fel, ha a feladatot a véges differenciák módszerével akarjuk megoldani. Ha a számunkra érdekes hullámhossz $\frac{2\pi}{k_c}$ -hez képest nagy, akkor elvileg az alap és σ -egyenletrendszert egyaránt használhatjuk, mert ebben a tartományban azonosan viselkednek.

RICHTMYER és MORTON [9], azonban rámutat, hogy a véges differenciák alkalmazásával a Δx rácstávolságnak megfelelő hullámhosszú perturbációt viszünk a rendszerbe. Biztosítani kell, hogy ennek növekedése ne legyen nagyobb, mint a vizsgált hullámhosszaké.

Mivel az alap-egyenletrendszerben (5.26) szerint minden hullámhossz instabil, ezért ha ezzel számolunk, akkor a modellbe kell foglalni valamilyen mesterséges stabilizálást a Δx nagyságrendű hullámhosszakra. Ilyen lehet pl. egy mesterséges felületi feszültség definiálása (5.23) alapján, ahol $k \cong \frac{\pi}{\Delta x}$ (ui. a rácson felvehető legkisebb hullámhossz $2\Delta x$):

$$(5.28) \quad \sigma^* = C \frac{(\Delta x)^2}{\pi^2 H} \left(\frac{\varrho_v \varrho_l}{(1-\alpha)\varrho_v + \alpha\varrho_l} \right) (v_v - v_l)^2,$$

ahol C egy dimenzió nélküli konstans, 1 körüli értékkel. Ez a σ^* általában jóval nagyobb mint a felületi feszültség fizikai értéke.

A σ -egyenletrendszer használatakor a Δx -nek megfelelő hullámhosszak stabilizálása az (5.23)-ban szereplő változók értékétől, valamint a Δx és σ nagyságától függ. A gyakorlatban előforduló tipikus értékek mellett ez a hullámhossz még instabil, a stabil hullámhosszak nagysága Δx -nél lényegesen kisebb hullámhosszaknál kezdődik. Ezért ebben az esetben is ajánlatos az (5.28) mesterséges felületi feszültség alkalmazása,

Természetesen nem ez az egyetlen módja a stabilizálásnak. Más fizikai hatások (viszkózitás, fázisátalakulás) figyelembe vétele is hasonló eredményre vezetnek.

Bizonyos differenciálási módszerek önmagukban is mesterséges stabilizáló hatással rendelkeznek anélkül, hogy külön stabilizáló tagot írtunk volna explicit módon az egyenletekbe. Ilyen pl. a 3) pontban leírt donor differenciálás [2].

A most ismertetett vizsgálat a felületi feszültség figyelembe vételén alapszik. Fizikailag még indokoltabb a viszkozitás figyelembe vétele mert az nemcsak stabilizáló, hanem csillapító hatással is rendelkezik. A vizsgálat rámutatott, hogy a kétfázisú áramlás két-közeg modelljét leíró egyenletek numerikus instabilitása megszüntethető, ha alkalmas fizikai jelenségeket beépítünk a modellbe. Ezekkel a karakterisztikák valósakká válnak és a rövid hullámhosszak stabilizálódnak.

IRODALOM

- [1] O'BRIEN, G. G., HYMAN, M. A. and KAPLAN, S., "A study of the numerical solution of partial differential equations", *Journal of Mathematics and Physics* **29** (1950) 223—240.
- [2] HARLOW, F. H. and AMSDEN, A. A., "A numerical fluid dynamics calculation method for all flow-speed", *Journal of Computational Physics* **8** (1971) 197—213.
- [3] HARLOW, F. H. and AMSDEN, A. A., "Numerical calculation of multiphase fluid flow", *Journal of Computational Physics* **17** (1975) 19—52.
- [4] HARLOW, F. H. and AMSDEN, A. A., "Flow of interpenetrating material phases", *Journal of Computational Physics* **18** (1975) 440—464.
- [5] LAX, P. D. and RICHTMYER, R. D., "Survey of the stability of linear finite difference equations", *Communications of Pure and Applied Mathematics* **9** (1956) 267—293.
- [6] LAX, P. D., "Differential equations, difference equations and matrix theory", *Communications of Pure and Applied Mathematics* **11** (1958) 187.
- [7] LILES, D. R. and REED, W. H., "A semi-implicit method for twophase fluid dynamics", *Journal of Computational Physics* **26** (1978) 390—407.
- [8] RAMSHAW, J. D. and TRAPP, J. A., "Characteristics, stability and short-wavelength phenomena in two-phase flow equation systems", *Nuclear Science and Engineering* **66** (1978) 93—102.
- [9] RICHTMYER, R. D. and MORTON, K. W., *Difference Methods for Initial Value Problems* (Interscience, New York, 1967).
- [10] STEWART, H. B., "Stability of two-phase flow calculation using two-fluid models", *Journal of Computational Physics* **33** (1979) 259—270.
- [11] WILKINSON, J. H., *The Algebraic Eigenvalue Problem* (University Press, Oxford, 1965).

(Beérkezett: 1982. december 17.)

DÚS MAGDOLNA
MTA KÖZPONTI FIZIKAI KUTATÓ INTÉZETE
1121 BUDAPEST, KONKOLY THEGE UT

MATHEMATICAL MODELLING OF TRANSIENT TWO-PHASE FLOWS

M. Dús

The paper investigates the stability of transient two-phase flow calculations using two-fluid models. Two-fluid modeling of two-phase flow may yield a system of partial differential equations having complex characteristics. This results in a mathematically ill-posed initial value problem.

Two-fluid models are solved generally by finite differences, but no finite difference method for solving such a problem can be stable in the usual sense. The paper investigates the numerical behavior of one such set of difference equation and derives conditions under which solutions appear to be well behaved.

On the other hand it shows that the relatively minor modification of introducing surface tension is sufficient to render the characteristics real and produce a well-posed problem.

A LAMBDA FÁG SZABÁLYOZÁSÁNAK MODELLEZÉSE

POLNER GÁBOR, DALLMANN GÉZA és OROSZ LÁSZLÓ

Szeged

A lambda fág az egyik legegyszerűbb és legjobban megismert élőlény. Szaporodása kétféleképpen történhet. Az, hogy a szaporodásnak melyik útja valósul meg, egy sztochasztikus szabályozási rendszer segítségével dől el. A dolgozat ezen szabályozási rendszer modellezésével foglalkozik.

Az analitikus vizsgálat egy *Kolgomorov-féle hátramutató differenciálegyenlet-rendszer* felállításához vezetett. A kapott egyenletrendszer bár elvileg megoldható, bonyolultsága miatt mégsem tartjuk célravezetőnek a megoldását.

A folyamat számítógépes szimulációját *Monte-Carlo módszerrel* végeztük el. Így lehetőségünk nyílt arra, hogy a működésre vonatkozó paraméterekből — melyeket kísérleti adatok alapján határoztak meg — rekonstruáljuk a lambda fág szabályozásának működését. Vizsgálatokat végeztünk, hogy néhány paraméter változtatására hogyan reagál a modell. Sajnos ellenőrzésül jóval kevesebb adatot tudunk felhasználni, mint ahány paraméter változtatására lehetőségünk van. Így várhatóan több olyan paraméter konfigurációt elő lehet állítani, melyek alapján a modell az ellenőrzési adatoknak megfelelően működik. Eredményeink ezt a várakozást támasztják alá. Előállítottunk néhány olyan paraméter konfigurációt, melyekkel az ellenőrzési adatoknak megfelelően működik a modell és a kísérleti megfigyelt értékektől sem különbözik lényegesen.

1. A lambda fág szabályozása

Az *Escherichia coli* bélbaktérium lambda fágja igen egyszerű élőlény. „Szervezete” csupán egy jellegzetes alakú fehérjeburokból áll, valamint a benne levő DNS molekulából. Tulajdonképpen vitatott, hogy élőlénynek tekinthető-e, ugyanis szaporodása egy sokkal fejlettebb élőlény — gazdabaktériuma — közreműködésével történik. A fág tehát önállóan szaporodni nem képes, viszont képes arra, hogy fágfertőzéskor gazdabaktériumához tapadva — a lambda fág esetén ez az *Escherichia coli* baktérium — a baktériumba juttassa DNS molekuláját. Ezután működésbe lép a bejutott DNS molekula által kódolt szabályozási rendszer. A fágtól csak a DNS molekula származik, a szabályozásban résztvevő összes többi alkotórész (az egész fehérjeszintetizáló apparátus) a baktérium szervezetéhez tartozik. A szabályozás eredményeképpen a lambda fág — és minden más ún. mérsékelt fág — esetében a szaporodásnak kétféle útja valósulhat meg: litikus és lizogén út.

A fág litikusan szaporodik, ha a baktériumba bejutó fág DNS molekula a baktériumot fágok felépítésére kényszeríti. Így a baktériumban megjelennek az utód fágok, melyek végül feloldják a gazdasejt sejtfalát. A gazdasejt elpusztul és a fágok kiszabadulnak, majd újabb baktériumokat fertőznek meg.

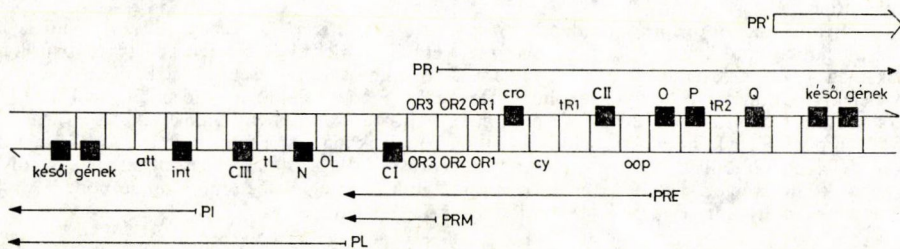
Lizogén szaporodás esetén a baktériumba bejutott fág DNS molekula beépül a baktérium DNS molekulájába. A baktérium osztódása során a baktérium DNS molekulája is megkettőződik, így mindegyik utód baktériumba jut DNS molekula. A baktérium DNS molekulájának részeként a fág DNS molekulája is jelen van az

utód baktériumban. Ekkor tulajdonképpen csak a fág DNS molekulája szaporodik. A gazdasejt nem károsodik, sőt újabb hasonló fággal történő fertőzés esetén a bejutó fág DNS molekula nem képes a litikus szaporodási útra térni, nem képes a baktériumot elpusztítani. A beépült fág DNS molekula helyzete viszonylag stabil, a tapasztalat szerint átlagosan 10^5 generációnként tud kiszakadni a baktérium DNS molekulájából, és a litikus szaporodási útra visszatérni.

A két szaporodási út közül választó szabályozási folyamat sztochasztikus folyamat, melynek eredményeként a tapasztalat szerint az esetek 30%-ában lizogén, 70%-ában litikus úton szaporodik a fág. A szabályozás végül is azon dől el, hogy — lizogén szaporodás esetén — a fág DNS molekula beépülését és a többi szabályozási folyamat gátlását előidéző, illetve — litikus szaporodás esetén — az utód fágok fehérjeburkát előállító folyamatok aktivizálódnak-e.

A dolgozatban a lambda fág litikus, illetve lizogén szaporodását eldöntő szabályozás modellezésével foglalkozunk.

Tekintsük át a szaporodás szabályozását az 1. ábra alapján!



1. ábra

Az ábra a lambda fág kettős szálú gyűrű alakú DNS molekulájának egy részletét ábrázolja kiterítve (a szabályozásban résztvevő szakaszokat a genetikában szokásos módon jelöltük [3]). A sötét téglalappal jelölt szakaszok az egyes gének. A többi megjelölt szakasz az att kivételével vezérlő funkciót lát el. Lizogén útra térésnél az att szakasznál nyílik szét a fág DNS molekula, és épül be a baktérium DNS molekulájába. A meg nem jelölt szakaszok a szabályozás szempontjából kevésbé lényeges biológiai funkciót töltenek be.

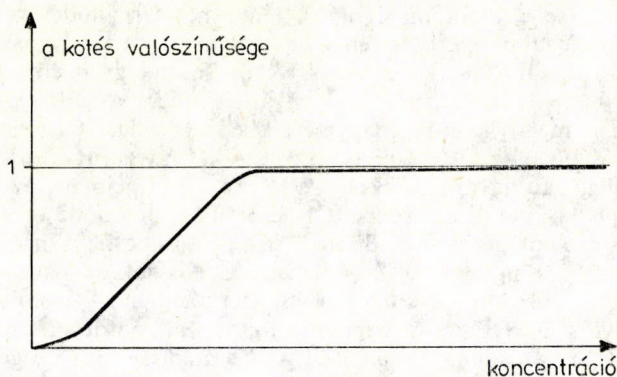
Egy gén megnyilvánul (az illető géntermék fehérje megtermelődik), ha a gén szakaszán transzkripció halad át. Transzkripciókor az RNS polimeráz enzim a DNS molekula egyik szála alapján ún. „hírvivő” RNS szálakat szintetizál, amely alapján a fehérjék szintézise történik. (A fehérjeszintézis leírása a gimnáziumi biológia tankönyvben megtalálható.) Egy hírvivő RNS molekula alapján csak néhány fehérje molekula termelődik, mert a hírvivő RNS molekula szintézise (a transzkripció) meghatározott módon zajlik. Csak a DNS molekula meghatározott helyeiről, az ún. promoterektől kiindulva a DNS molekula meghatározott szála alapján meghatározott irányban meghatározott hosszúságban történhet a transzkripció. A lambda fág esetében a már kész fehérje molekulák is befolyással vannak a további transzkripciókra. Van olyan transzkripció, amely csak akkor tud elindulni, ill. indulásának valószínűsége növekszik, ha promoteréhez meghatározott fehérje kötődik. (A kötés természetét később tárgyaljuk.) A vezérlő szakaszok kétféleképpen lehetnek. Az egyik típusún csak akkor tud a transz-

kripció végighaladni, ha ahhoz a megfelelő fehérjék kötődnek, a másik típusún viszont csak akkor tud a transzkripció végighaladni, ha ahhoz nem kötődik fehérje.

Az 1. ábra két oldalán levő nyilak mentén haladhatnak a transzkripciók. Az ábra azon oldalához vannak csoportosítva, amely DNS szálon haladnak. A PR' transzkripciót azért jelöltük vastagabb nyíllal mint a többi, mert végbemenetele esetén körülbelül tízszer több hírvivő RNS molekulát szintetizál mint a PR.

Az egyes géntermék fehérjéknél a DNS molekulához való kötésének valószínűsége az illető fehérje koncentrációjától függ. Alacsony koncentráció esetén a kötés valószínűsége megközelítően egyenesen arányos a koncentrációval. A kötés természetét a 2. ábrán látható függvény írja le. A görbéről láthatjuk, hogy van egy küszöb koncentráció érték, melytől kezdve a kötés valószínűsége megközelítően egy. Ettől a küszöbtől kezdve a kötést tekinthetjük „tartósnak”. Valójában a már egyszer létrejött kötés korántsem tekinthető stabilnak. A fehérje molekula koncentráció függvényében a kötés valószínűségét az előbb említett függvény írja le.

Ezek után tekintsük át a szabályozást részletesen!



2. ábra

Először a PL és a PR transzkripciók mennek végbe a tL, ill. a tR1 szakaszokig [5].

A PL transzkripció során megnyilvánul az N gén, melynek terméke kötődik a tL, tR1, tR2 szakaszokhoz. Az a hatása, hogy az ezeken a szakaszokon áthaladó — különben megálló — PL, PR transzkripciókat továbbengedi [5].

A PR transzkripció során megnyilvánul a cro gén, melynek terméke kötődik az OL és OR3, OR2, OR1 szakaszokhoz. A cro gén terméke az OR3, OR2, OR1 szakaszokhoz különböző erősséggel kötődik: legerősebben az OR3-hoz, leggyengébben az OR1-hez. Hatása: az OL, OR3, OR2, OR1 szakaszokon mindkét irányból áthaladó transzkripciókat nem engedi tovább [1].

A tL és a tR1 szakaszokon túljutva megnyilvánulnak a CII és a CIII gének, melyek termékei együtt kötődnek az oop és a cy szakaszokhoz, lehetővé téve a PRE transzkripció végighaladását [5].

A CII és a CIII géntermékek együtt kötődnek még az int gén mellé, lehetővé téve annak közvetlen PI transzkripcióját. Az int gén terméke a fág DNS integrálódásához (beépülés a baktérium DNS-ébe) szükséges. Bár int transzkripciója a PL transzkripció

során is megvalósulhat, az integrálódáshoz szükséges int géntermék döntő többsége int közvetlen PI transzkripciójából származik [6].

A PR transzkripció során a tR2 szakasz eléréséig az O és a P gének nyilvánulnak meg, melyek géntermékei a fág DNS sokszorozódásához szükségesek. O és P triviálisan szükséges a litikus szaporodáshoz. O és P azonban a lizogén szaporodáshoz is szükséges, ugyanis O és P mutáns fágok (O-ban és P-ben sérült fágok) nem képesek lizogén úton szaporodni [7].

A PRE transzkripció során nyilvánul meg a CI gén, melynek génterméke az ún. represszor molekula az OL, OR3, OR2, OR1 szakaszokhoz kötődik. A CI géntermék sokkal erősebben kötődik ezekhez a szakaszokhoz, mint a cro géntermék. A CI géntermék az OR3, OR2, OR1 szakaszokhoz különböző erősséggel kötődik: legerősebben OR1-hez, leggyengébben OR3-hoz. A CI gén termékének hatása hasonló a cro-éhoz: az OL, OR3, OR2, OR1 szakaszokon mindkét irányból áthaladó transzkripciókat nem engedi tovább. A cro és a CI géntermékek antagonisztikus volta az OR szakaszokhoz való különböző természetű kötődésük alapján magyarázható. Ezek a géntermékek ugyanis úgy kötődnek az OR szakaszokhoz, hogy saját termelődésüket serkentik vagy csak kissé gátolják, míg a másik géntermék termelődését nagymértékben akadályozzák. Pontosabban, cro esetében cro géntermék legerősebben az OR3 szakaszhoz kötődik, az OR2 és OR1 szakaszokhoz sokkal gyengébben. Tehát a PR transzkripciót, mely során cro megnyilvánul alig, a PRE transzkripciót, mely során CI megnyilvánul, nagymértékben gátolja. CI esetében a CI géntermék legerősebben az OR1 szakaszhoz kötődik, így a PR transzkripciót nagymértékben gátolja, gátolja ugyanakkor a PRE transzkripciót is, azonban a PRM transzkripció egyidejű aktivizálásával a CI géntermék termelődése végeredményben növekszik. CI és cro antagonizmusa döntő a szaporodás későbbi litikus vagy lizogén volta szempontjából is, ugyanis a szabályozás kezdeti szakaszában a cro magasabb koncentrációja a litikus, míg CI magasabb koncentrációja a lizogén szaporodás választásának kedvez. A PRE transzkripciót a lizogén állapotot létrehozó transzkripciónak is nevezik, ugyanis hatására az OR1, OR2 szakaszokhoz CI géntermék kötődik, így a PR transzkripció a továbbiakban nem lehetséges (ami szükséges lenne a litikus útra téréshez), ugyanakkor lehetővé teszi a CI gén további megnyilvánulását a PRM transzkripció aktivizálásával [1], [2].

CI további megnyilvánulásával a PRM transzkripció során az OL, OR3 szakaszokhoz CI géntermék kötődik, így csupán az int gén megnyilvánulása lehetséges a PI transzkripció során. A fág DNS molekula szétnyílik az att régiójánál és integrálódik a baktérium DNS molekulájába annak att régiójánál [5]. Ezzel a szaporodás milyensége eldőlt, a fág lizogén úton szaporodik.

A lizogén állapotban a CI géntermék bomlásával annak koncentrációja csökken, de az OR3, OR2, OR1 szakaszokhoz való kötődés alapján először az OR3 szakaszból disszociál le — az OR1, OR2 szakaszokon még maradván annyi CI termék, hogy a PR transzkripció során cro ne nyilvánulhasson meg —, egyből lehetővé téve, hogy a lizogén állapotot fenntartó PRM transzkripció során a CI gén megnyilvánuljon.

Ha a PR transzkripció túljut a tR2 szakaszon, a Q gén nyilvánul meg, melynek terméke a késői gének igen gyors megnyilvánulását eredményezi a PR' transzkripció során [5]. A késői gének termékei a fág fehérjeburkához szükségesek. A gyűrű alakú DNS molekulában a késői gének egy része a PR és PR' transzkripciók során, másik része a PL transzkripció során nyilvánul meg. Érdekes, hogy van a késői

géneknek egy olyan DNS szakasza, melynek egyik szálán a PR és PR', másik szálán a PL transzkripció halad végig [3]. A PR és PR' valamint PL transzkripciók során megtörténik a már korábban az O és P géntermékek közreműködésével megsztintetizált fág DNS molekulák és a késői gének által kódolt fehérjeburok alkotórészek összeállása fágokká. Ezzel eldől, hogy a fág a litikus szaporodás útjára lépett.

2. Egyszerűsítések a lambda fág szabályozásában

A modellezés során az előbbieken ismertetett szabályozási modellt használtuk fel.

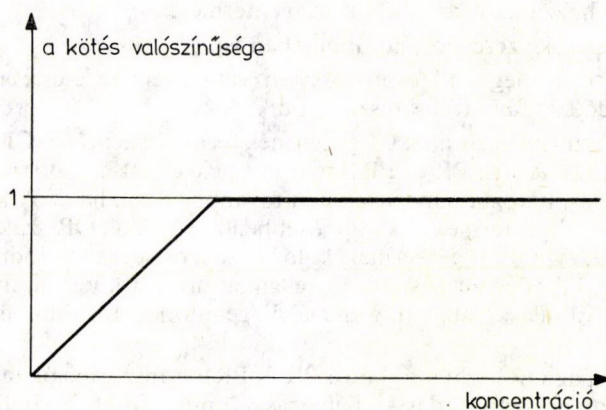
A modellen bizonyos egyszerűsítéseket végeztünk. A megvalósított modell a 4. ábra alapján követhető nyomon.

a) A fág DNS beépülésével nem foglalkoztunk, mivel a lizogén állapot létrehozása és a beépülés koordináltan megy végbe. Ugyanis a lizogén állapot létrehozásához CI megnyilvánulása szükséges a PRE transzkripció során, a beépüléshez int megnyilvánulása a PI transzkripció során. Mindkét transzkripció csak kellő CII és CIII géntermék koncentráció során jöhet létre. (Az olyan esetek igen ritkák, amikor a CI géntermék tartósan kötődik az OL, OR3, OR2, OR1 szakaszokhoz, de beépülés nem következik be.)

b) Nem foglalkoztunk a fág DNS replikációjához szükséges O és P génekkel, mert ezek mind a litikus mind a lizogén szaporodáshoz szükségesek és meg tudnak nyilvánulni abban a stádiumban, amikor a szaporodás milyensége még nem dől el.

c) A PRM transzkripciót nem vettük figyelembe, ugyanis ez a transzkripció csak a lizogén állapot fenntartásában, nem pedig létrehozásában játszik szerepet.

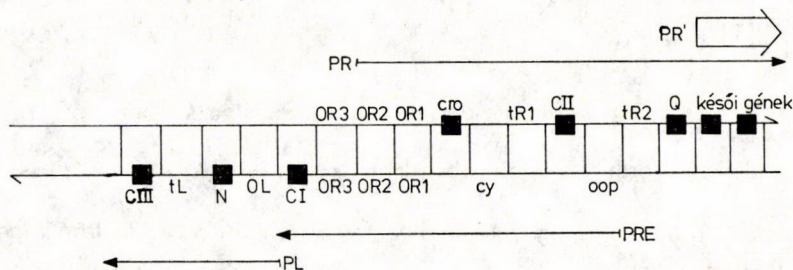
d) A számítógépes szimuláció során a géntermék kötődését a DNS molekula bizonyos részeihez nem a 2. ábrán látható menetű függvény alapján, hanem ennek közelítéseként a 3. ábrán megadott menetű függvény alapján számoltuk.



3. ábra

3. A modellezés alapjául szolgáló egyszerűsített lambda fág szabályozás

A leírtak a 4. ábra alapján követhetők nyomon. Az ábra felépítése hasonló az 1. ábrához.



4. ábra

A regulációban a PL, PR, PRE transzkripciók vesznek részt, melyek véletlenül különböző valószínűséggel indulnak el a transzkripciók kezdőpontjából. Az egyes transzkripciók során olyan termékek szintetizálódnak, melyek szabályozzák a transzkripciók haladását, azaz további termékek szintetizálódását.

A PL transzkripció során az N és a CII géntermékek szintetizálódhatnak meg.

Az N géntermék a tL, tR1, tR2 szakaszokhoz kötődik. Hatására a PL transzkripció a tL szakaszon tovább tud haladni. Hasonlóan a PR transzkripció a tR1 és a tR2 szakaszokon tovább tud haladni.

A CII géntermék a oop és cy szakaszokhoz kötődik a CIII géntermékkel együtt. Hatásuk: a PRE transzkripció csak akkor tud az oop és cy szakaszokon tovább haladni, ha azokhoz CII és CIII géntermékek együtt kötődnek.

A PR transzkripció során a cro, CII és Q géntermékek szintetizálódhatnak meg.

A cro géntermék az OL, OR3, OR2, OR1 szakaszokhoz kötődik. Hatása: a PL transzkripció csak akkor tud OL-en továbbhaladni, ha OL-hez nem kötődik cro géntermék; hasonlóan a PR transzkripció OR2-n és OR1-en csak akkor tud továbbhaladni, ha hozzájuk nem kötődik cro géntermék.

A CII géntermék szerepét a korábbiakban ismertettük.

A Q géntermék elegendő mennyiségben való jelenléte esetében a szaporodás szabályozása eldőlt: a fág litikusan szaporodik.

A PRE transzkripció során a CI géntermék termelődhet meg. CI termék hasonló a cro géntermékhez. A PL, PR és PRE transzkripciók csak akkor tudnak OL, OR3, OR2, OR1 szakaszok közül a megfelelőkön továbbhaladni, ha azokhoz nem kötődik CI géntermék. A CI géntermék sokkal erősebben kötődik az OR szakaszokhoz, mint a cro géntermék. Mindkét géntermék különböző erősséggel kötődik az egyes OR szakaszokhoz. E kötődés természetét az jellemzi, hogy a kötés által cro és CI saját további megnyilvánulását alig, míg a másik géntermék további megnyilvánulását erősen gátolja.

Amennyiben az OR1 szakaszokon a PR és PRE transzkripciók tartósan nem tudnak továbbhaladni, a szaporodás szabályozása szintén eldőlt: a fág lizogén úton fog szaporodni.

4. Analitikus modellezés differenciálegyenletrendszerrel

A lambda fág szabályozásának leírására *Markov folyamatot* használunk, melynek állapotait az egyes gének megnyilvánulásainak számai jelzik.

A *Markov folyamatunk* állapotai a következők.

$$\xi_{i,j;k,l,m,n;o}^{\text{CIII}, \text{N}; \text{késői}, \text{Q}, \text{CII}, \text{cro}; \text{CI}}$$

Ahol $\xi_{i,j;k,l,m,n;o}^{\text{X}}$ jelzi, hogy X géntermék y -szor termelődött meg. Az i, j, k, l, m, n, o indexek nem negatív egész értékeket vehetnek fel, csupán az a kikötés, hogy a lambda fág szabályozásával elérhető állapot legyen. Ugyanis a géntermékek megtermelődése egymástól nem független: ha bizonyos géntermék megtermelődik, más géntermékeknek is meg kell termelődniük. Ezt a problémát az indexértékek összegekre bontásával fogjuk megoldani, ahol az egyes összeadandók már tetszőleges nemnegatív értéket felvehetnek.

Az N termék termelődésére a 3. pontban leírtak alapján a következő lokális egyenletek írhatók fel:

$$(4.1) \quad P(\xi_{i+h}^{\text{N}} = i+1 | \xi_i^{\text{N}} = i, \xi_i^{\text{CI}} = k, \xi_i^{\text{cro}} = l) = \lambda_{kl}^{\text{N}} h + o(h)$$

$$(4.2) \quad P(\xi_{i+h}^{\text{N}} = i | \xi_i^{\text{N}} = i, \xi_i^{\text{CI}} = k, \xi_i^{\text{cro}} = l) = 1 - \lambda_{kl}^{\text{N}} h + o(h)$$

$$(4.3) \quad P(\xi_{i+h}^{\text{N}} \geq i+2 | \xi_i^{\text{N}} = i) = o(h)$$

Ahol $\lambda_{kl}^{\text{N}} = \lambda^{\text{PL}}(1 - f^{\text{CI}, \text{OL}}(k + x^{\text{OL}}l))$ és $k = 0, 1, 2, \dots$ és $l = 0, 1, 2, \dots$

Az egyes változók jelentése:

ξ_i^{X} jelöli X géntermék hányszor nyilvánult meg t idő alatt.

λ^{PL} jelöli annak a folyamatnak a paraméterét, amellyel a PL transzkripció elindulhat.

$f^{\text{CI}, \text{OL}}(x)$ jelöli azt a függvényt, mely CI géntermék OL-hez való kötődésének valószínűségét adja meg.

Mivel az OL szakaszhoz CI és cro egyaránt kötődik, x^{OL} megmutatja, hogy 1 megnyilvánulásnyi cro géntermék x^{OL} megnyilvánulásnyi CI géntermékkel egyenértékűen tud az OL-hez kötődni.

Így annak a valószínűsége, hogy az OL szakaszhoz CI vagy cro kötődik $f^{\text{CI}, \text{OL}}(k + x^{\text{OL}}l)$. Tehát az N termék termelődésének a λ_{kl}^{N} paraméterére $\lambda_{kl}^{\text{N}} = \lambda^{\text{PL}}(1 - f^{\text{CI}, \text{OL}}(k + x^{\text{OL}}l))$ adódik, ugyanis N megnyilvánulásához a PL transzkripciónak át kell haladni az OL szakaszon.

k és l miatt első ránézésre végtelen sok lokális egyenletünk van, azonban tekintettel arra, hogy egy sejtben játszódik le a folyamat, cro és CI nem nyilvánulhat meg tetszőlegesen sokszor, így k és l gyakorlatilag nem végtelenig, hanem egy elég nagy K értékig vehet fel értékeket.

Az N termék képződéséhez hasonlóan a CIII termék képződésére az alábbi lokális egyenletek írhatók fel $k=0, 1, 2, \dots$, $l=0, 1, 2, \dots$ és $m=0, 1, 2, \dots$ esetére:

$$(4.4) \quad P(\xi_{i+h}^{\text{CIII}} = i+1 | \xi_i^{\text{CIII}} = i, \xi_i^{\text{CI}} = k, \xi_i^{\text{cro}} = l, \xi_i^{\text{N}} = m) = \lambda_{klm}^{\text{CIII}} h + o(h)$$

$$(4.5) \quad P(\xi_{i+h}^{\text{CIII}} = i | \xi_i^{\text{CIII}} = i, \xi_i^{\text{CI}} = k, \xi_i^{\text{cro}} = l, \xi_i^{\text{N}} = m) = 1 - \lambda_{klm}^{\text{CIII}} h + o(h)$$

$$(4.6) \quad P(\xi_{i+h}^{\text{CIII}} \geq i+2 | \xi_i^{\text{CIII}} = i) = o(h),$$

ahol

$$\lambda_{klm}^{\text{CII}} = \lambda^{\text{PL}}(1 - f^{\text{CI,OL}}(k + \kappa^{\text{OL}}l))f^{\text{N,IL}}(m).$$

Hasonlóan az eddigiekhez a cro termék képződésére az alábbi lokális egyenletek írhatók fel $k=0, 1, 2, \dots$ esetére:

$$(4.7) \quad P(\xi_{t+h}^{\text{cro}} = i + 1 | \xi_t^{\text{cro}} = i, \xi_t^{\text{CI}} = k) = \lambda_{ik}^{\text{cro}}h + o(h)$$

$$(4.8) \quad P(\xi_{t+h}^{\text{cro}} = i | \xi_t^{\text{cro}} = i, \xi_t^{\text{CI}} = k) = 1 - \lambda_{ik}^{\text{cro}}h + o(h)$$

$$(4.9) \quad P(\xi_{t+h}^{\text{cro}} \geq i + 2 | \xi_t^{\text{cro}} = i) = o(h),$$

ahol

$$\lambda_{ik}^{\text{cro}} = \lambda^{\text{PL}}(1 - f^{\text{CI,OR2}}(k + \kappa^{\text{OR2}}i))(1 - f^{\text{CI,OR1}}(k + \kappa^{\text{OR1}}i)).$$

A CII termék képződésére az alábbi lokális egyenletek írhatók fel $k=0, 1, 2, \dots$, $l=0, 1, 2, \dots$ és $m=0, 1, 2, \dots$ esetére:

$$(4.10) \quad P(\xi_{t+h}^{\text{CII}} = i + 1 | \xi_t^{\text{CII}} = i, \xi_t^{\text{CI}} = k, \xi_t^{\text{cro}} = l, \xi_t^{\text{N}} = m) = \lambda_{klm}^{\text{CII}}h + o(h)$$

$$(4.11) \quad P(\xi_{t+h}^{\text{CII}} = i | \xi_t^{\text{CII}} = i, \xi_t^{\text{CI}} = k, \xi_t^{\text{cro}} = l, \xi_t^{\text{N}} = m) = 1 - \lambda_{klm}^{\text{CII}}h + o(h)$$

$$(4.12) \quad P(\xi_{t+h}^{\text{CII}} \geq i + 2 | \xi_t^{\text{CII}} = i) = o(h),$$

ahol

$$\lambda_{klm}^{\text{CII}} = \lambda^{\text{PR}}(1 - f^{\text{CI,OR2}}(k + \kappa^{\text{OR2}}l))(1 - f^{\text{CI,OR1}}(k + \kappa^{\text{OR1}}l))f^{\text{N,IR1}}(m).$$

A CI termék képződése az alábbi lokális egyenletek szerint történik $k=0, 1, 2, \dots$, $l=0, 1, 2, \dots$ és $m=0, 1, 2, \dots$ esetére:

$$(4.13) \quad P(\xi_{t+h}^{\text{CI}} = i + 1 | \xi_t^{\text{CI}} = i, \xi_t^{\text{cro}} = k, \xi_t^{\text{CII}} = l, \xi_t^{\text{CIII}} = m) = \lambda_{iklm}^{\text{CI}}h + o(h)$$

$$(4.14) \quad P(\xi_{t+h}^{\text{CI}} = i | \xi_t^{\text{CI}} = i, \xi_t^{\text{cro}} = k, \xi_t^{\text{CII}} = l, \xi_t^{\text{CIII}} = m) = 1 - \lambda_{iklm}^{\text{CI}}h + o(h)$$

$$(4.15) \quad P(\xi_{t+h}^{\text{CI}} \geq i + 2 | \xi_t^{\text{CI}} = i) = o(h),$$

ahol

$$\lambda_{iklm}^{\text{CI}} = \lambda^{\text{PRE}}f^{\text{CII,oop}}(l)f^{\text{CIII,oop}}(m)f^{\text{CII,cy}}(l)f^{\text{CIII,cy}}(m) \times \\ \times (1 - f^{\text{CI,OR1}}(i + \kappa^{\text{OR1}}k))(1 - f^{\text{CI,OR2}}(i + \kappa^{\text{OR2}}k))(1 - f^{\text{CI,OR3}}(i + \kappa^{\text{OR3}}k)).$$

A Q termék képződésére az alábbi lokális egyenletek írhatók fel $k=0, 1, 2, \dots$, $l=0, 1, 2, \dots$ és $m=0, 1, 2, \dots$ esetére:

$$(4.16) \quad P(\xi_{t+h}^{\text{Q}} = i + 1 | \xi_t^{\text{Q}} = i, \xi_t^{\text{CI}} = k, \xi_t^{\text{cro}} = l, \xi_t^{\text{N}} = m) = \lambda_{klm}^{\text{Q}}h + o(h)$$

$$(4.17) \quad P(\xi_{t+h}^{\text{Q}} = i | \xi_t^{\text{Q}} = i, \xi_t^{\text{CI}} = k, \xi_t^{\text{cro}} = l, \xi_t^{\text{N}} = m) = 1 - \lambda_{klm}^{\text{Q}}h + o(h)$$

$$(4.18) \quad P(\xi_{t+h}^{\text{Q}} \geq i + 2 | \xi_t^{\text{Q}} = i) = o(h),$$

ahol

$$\lambda_{klm}^{\text{Q}} = \lambda^{\text{PR}}(1 - f^{\text{CI,OR2}}(k + \kappa^{\text{OR2}}l))(1 - f^{\text{CI,OR1}}(k + \kappa^{\text{OR1}}l))f^{\text{N,IR1}}(m)f^{\text{N,IR2}}(m).$$

A jobb oldali késői gének a PR és a PR' transzkripciók során nyilvánulhatnak meg. A következő lokális egyenletek írhatók fel $j=0, 1, 2, \dots$, $k=0, 1, 2, \dots$, $l=0, 1, 2, \dots$ és $m=0, 1, 2, \dots$ esetére:

$$(4.19) \quad P(\xi_{t+h}^{\text{késői}} = i+1 | \xi_t^{\text{késői}} = i, \xi_t^{\text{CI}} = j, \xi_t^{\text{cro}} = k, \xi_t^{\text{N}} = l, \xi_t^{\text{Q}} = m) = \\ = \lambda_{ji}^{\text{K1}} h + \lambda_{mi}^{\text{K2}} h + o(h)$$

$$(4.20) \quad P(\xi_{t+h}^{\text{késői}} = i | \xi_t^{\text{késői}} = i, \xi_t^{\text{CI}} = j, \xi_t^{\text{cro}} = k, \xi_t^{\text{N}} = l, \xi_t^{\text{Q}} = m) = \\ = 1 - \lambda_{ji}^{\text{K1}} h - \lambda_{mi}^{\text{K2}} h + o(h)$$

$$(4.21) \quad P(\xi_{t+h}^{\text{késői}} \geq i+2 | \xi_t^{\text{késői}} = i) = o(h),$$

ahol

$$\lambda_{ji}^{\text{K1}} = \lambda_{jkl}^{\text{Q}} = \lambda^{\text{PR}} (1 - f^{\text{CI, OR2}}(j + \kappa^{\text{OR2}} k)) (1 - f^{\text{CI, OR1}}(j + \kappa^{\text{OR1}} k)) \times \\ \times f^{\text{N, tR1}}(l) f^{\text{N, tR2}}(l); \text{ illetve } \lambda_{mi}^{\text{K2}} = \lambda^{\text{PR'}} f^{\text{Q}}(m).$$

Most már fel tudjuk írni a lambda fág szabályozásának egyenletét, amely egy *Markov folyamat*, melynek állapotait az egyes gének megnyilvánulásainak számai jelzik. A *Kolmogorov hátramutató egyenletrendszer* írjuk fel.

A felírható egyenletrendszer az 5. ábrán látható. A folyamat tiszta születési folyamat, ugyanis ha egy géntermék megtermelődik, ő maga, vagy vele együtt termelődik olyan termék, mely a géntermék további termelődését szabályozza. Az egyenletrendszer a géntermékek változását írja le, azonban ez az egyenletrendszer magába foglalja azokat az egyenleteket is, amikor a géntermékek nem változnak. Ehhez azokat az átmenetvalószínűségeket, melyek valamely X géntermék változását i -ről j -re írják le, és $i > j$, célszerűen nullának definiáljuk. Azaz legyen

$$P_{i,j;\dots}^{X;\dots}(t) = 0; \quad i > j.$$

Az 5. ábra egyenletrendszerét rendezve, felhasználva a (4.1)–(4.21) lokális egyenleteket, valamint a $h \rightarrow 0$ határmenetet képezve a 6. ábrán látható egyenletrendszerhez jutunk.

A 6. ábra egyenletrendszere, bár elég összetett, analitikusan megoldható. A rekurzív megoldás lépései a következők:

1. Egyik géntermék sem változik, azaz $j=0$, $l=0$, $o=0$, $p=0$, $r=0$, $v=0$, $x=0$ esetére meghatározzuk az átmenet valószínűségeket.

2. Rendre meghatározzuk azokat az átmenet valószínűségeket, amikor a j , l , o , p , r , v , x indexek közül egy, kettő és így tovább hét index nem nulla.

Ez a megoldás, bár elvi akadályai nincsenek, elég körülményesnek ígérkezik és minden bizonnyal nehéz numerikus matematikai problémákat rejt magába. A következő fejezetben a megoldásnak egy másik módjával a számítógépes szimulációval fogunk foglalkozni.

$$\begin{aligned}
 & i, j, \dots, x = 0, 1, 2, \dots; \quad t \equiv 0 \\
 & P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} = \\
 & = -P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times (\lambda_{w+x,m+q+s+o+r+v}^N + \lambda_{m+q+s+o+r+v,w+x}^{\text{cro}} + \lambda_{w+x,m+q+s+o+r+v,m+q+s+o+r+v}^{\text{CI}} \times \lambda_{w+x,m+q+s+o+r+v}^N) + \\
 & + P_{i,i+j;i+k,i+k+j+l-1;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \lambda_{w+x,m+q+s+o+r+v}^N \\
 & + P_{i,i+j-1;i+k,i+k+j+l-1;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times \lambda_{w+x,m+q+s+o+r+v,i+k+j+l-1}^{\text{CIII}} + \\
 & \delta/a. \text{ ábra}
 \end{aligned}$$

$$\begin{aligned}
 & + P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p-1;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \lambda_{m+o}^{\text{K}_2} + \\
 & + P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v-1;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times \lambda_{m+q+s+o+r+v-1,w+x}^{\text{cro}} + \\
 & + P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r-1;m+q+s,m+q+s+o+r+v;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times \lambda_{w+x,m+q+s+o+r+v-1,i+k+j+l}^{\text{CII}} + \\
 & + P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p-1;m,m+o-1;m+q,m+q+o+r-1;m+q+s,m+q+s+o+r+v-1;w,w+x}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times \lambda_{w+x,m+q+s+o+r+v-1,i+k+j+l}^Q + \\
 & + P_{i,i+j;i+k,i+k+j+l;m+n,m+n+o+p;m,m+o;m+q,m+q+o+r;m+q+s,m+q+s+o+r+v;w,w+x-1}^{\text{CIII}, N; \text{késői}, Q, \text{CII}, \text{cro}; \text{CI}} \times \\
 & \times \lambda_{w+x-1,m+q+s+o+r+v,m+q+o+r,i+j}^{\text{CI}} \\
 & \delta/b. \text{ ábra}
 \end{aligned}$$

Ha mégis birtokunkban lennének az átmenetvalószínűség értékek, meg tudnánk adni a lizogén út választásának valószínűségét.

$$P_{\text{lizogén}} = \lim_{t \rightarrow \infty} \sum_{x=S}^{\infty} \sum_{v=0}^{\infty} \sum_{r=0}^{\infty} \sum_{P=0}^L \sum_{o=0}^{L-P} \sum_{p=0}^P \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \times \\ \times P_{0,j;0,j+l;0,o+p;0,o;0,o+r,0,o+r+v;0,x}^{\text{CIH, N; késői; Q; CIH, cro; CI}}(t).$$

Ahol S jelöli a CI gén szükséges megnyilvánulásainak számát a lizogén útra téréshez. A késői génnek L -szeri megnyilvánulása pedig még nem jelent litikus fejlődést, $L+1$ -szeri megnyilvánulása pedig igen. A képletben szereplő végtelen összegzések tulajdonképpen praktikusán csak véges összegek, ugyanis a folyamat egy sejten belül játszódik le, így egy-egy gén csak véges sokszor nyilvánulhat meg. A $t \rightarrow \infty$ határmenet számolásakor is elegendő egy elég nagy t értéket figyelembe venni, ugyanis tapasztalat szerint a lambda fág szaporodásának szabályozása a fágfertőzést követően körülbelül fél óra alatt eldől.

5. A számítógépes szimuláció

A szimulációt *Monte-Carlo módszerrel* végeztük. A 3. pontban leírt modellt használtuk fel a folyamat modellezésére.

Véletlenszám generátorral szimuláltuk az egyes transzkripciók elindulását és a transzkripcióknak az OL, OR3, OR2, OR1, tL, tR1, tR2, oop, cy szakaszokon való továbbhaladását.

Véletlenszám generátornak a következőt használtuk [4]:

$$(5.1) \quad V_n = (2^7 + 1)V_{n-1} + 1 \pmod{2^{35}}$$

$$(5.2) \quad X_n = V_n \cdot 38/2^{35},$$

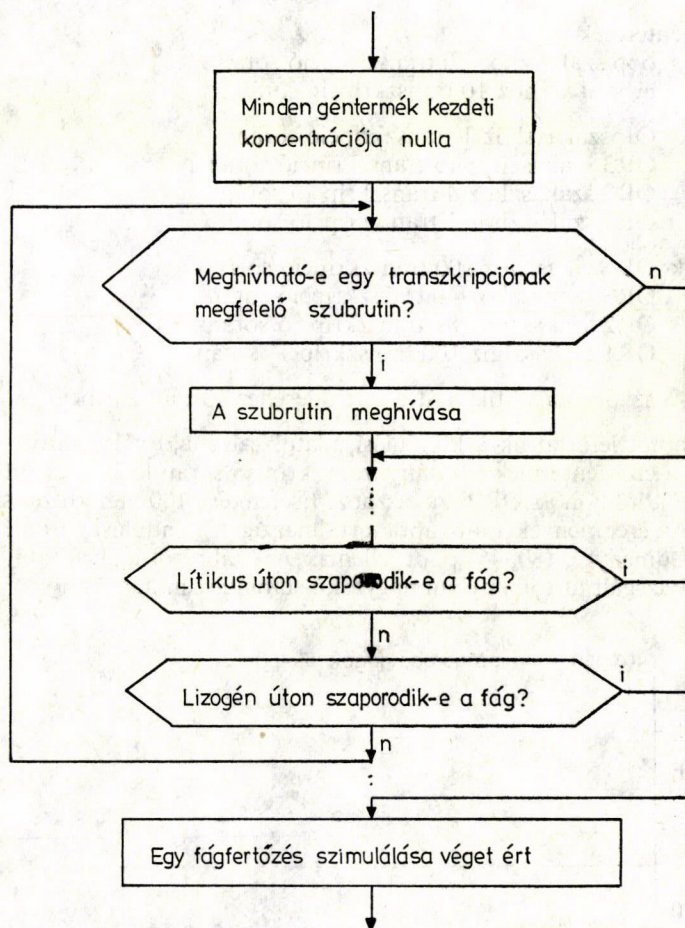
ahol X_n az n -edik véletlenszám, V_0 értéke tetszőlegesen választható 0-tól $2^{35} - 1$ -ig. Így az (5.1), (5.2) egyenletekkel leírt véletlenszám generátor által szolgáltatott véletlenszámok sorozatának a periódusa 2^{35} [4].

A program a 7. ábrán látható blokksema alapján egymás után meghívja az egyes transzkripcióknak megfelelő szubrutinokat az illető transzkripció indulási valószínűségének megfelelően. Ezek a szubrutinok a véletlenszám generátor alapján eljutnak egy bizonyos pontig az egyes géntermékek megszintetizálásával. A főprogram rendszeresen ellenőrzi, hogy a szaporodás litikus vagy lizogén módja eldőlt-e már, amennyiben igen, úgy leállítja a fágfertőzés szimulálását és újabb fágfertőzés szimulálásába kezd.

A programot PL/1 nyelven írtuk, és a JATE Kalmár László Kibernetikai Laboratóriumának R-40 számítógépén futtattuk OS rendszerben.

6. Futási eredmények

A lambda fág szabályozási modelljének paraméterei biokémiai becslésekből származnak. A számítógépes program segítségével vizsgálhatjuk, hogy a paraméterekkel lefuttatva a programot, mennyiben felel meg a természetben megfigyelt lambda fágoknak. Egyúttal ellenőrizhetjük azt is, hogy a lambda fág modellje mennyire írja le



7. ábra

a lambda fágot. Ellenőrzésül sajnos csupán a 30—70%-os lizogén-litikus szaporodási arányt tudjuk felhasználni.

A paraméterek a következők:

Az egyes transzkripciók indulási valószínűségeinek aránya:

$$P(PR):P(PL):P(PRE) = 13,5:6,5:13,5.$$

Az egyes géntermékeknek hány transzkripció során kell megszintetizálódniuk, hogy a megfelelő szakaszokhoz tartósan kötődjenek:

N géntermék tL szakaszokhoz 2 transzkripció során
 tR1 szakaszokhoz 3 transzkripció során
 tR2 szakaszokhoz 15 transzkripció során

CII és CIII géntermék

oop szakaszhoz 10 transzkripció során
cy szakaszhoz 10 transzkripció során

CI géntermék

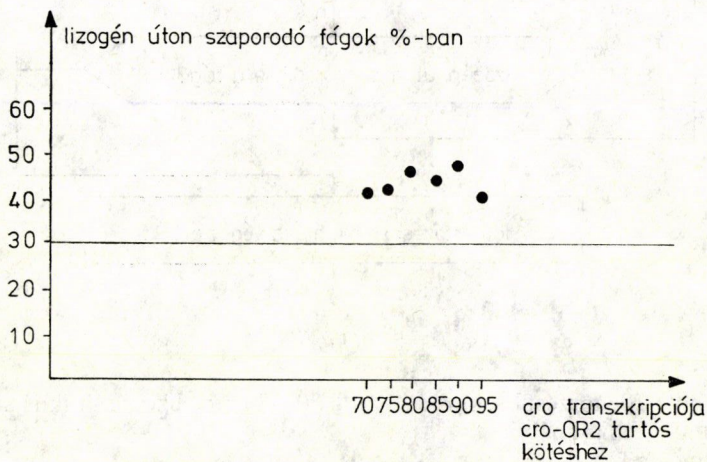
OL szakaszhoz 1 transzkripció során
OR3 szakaszhoz 15 transzkripció során
OR2 szakaszhoz 4 transzkripció során
OR1 szakaszhoz 1 transzkripció során

cro géntermék

OL szakaszhoz 100 transzkripció során
OR3 szakaszhoz 40 transzkripció során
OR2 szakaszhoz 80 transzkripció során
ORI szakaszhoz 100 transzkripció során

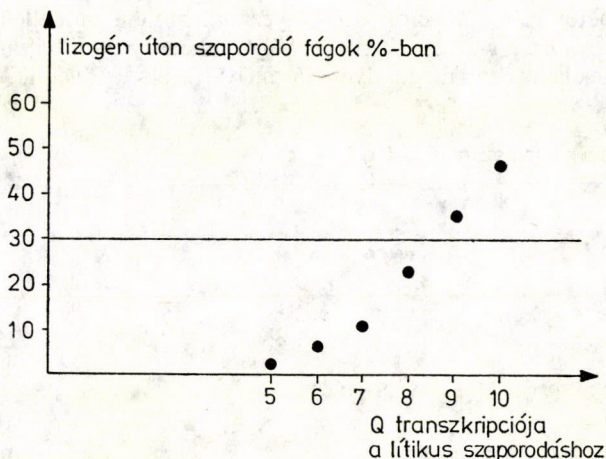
Q géntermék a szaporodás litikus útra terelődéséhez 10 transzkripció során.

A programot lefuttattuk a kiindulási adatokkal, majd változtattuk azt a paramétert, hogy a cro génterméknek hány transzkripció során kell megszintetizálódnia, hogy tartósan kötődjön az OR2 szakaszhoz. Esetenként 100 fágfertőzés szimulálását végeztük el. Az eredményeket a 8. ábra tartalmazza. A kiindulási paraméter értékekkel kapott eredményt a (80, 46) pont jellemzi. A 8. ábráról leolvasható, hogy a fág viselkedése a vizsgált tartományban nagyfokú stabilitást mutat.



8. ábra

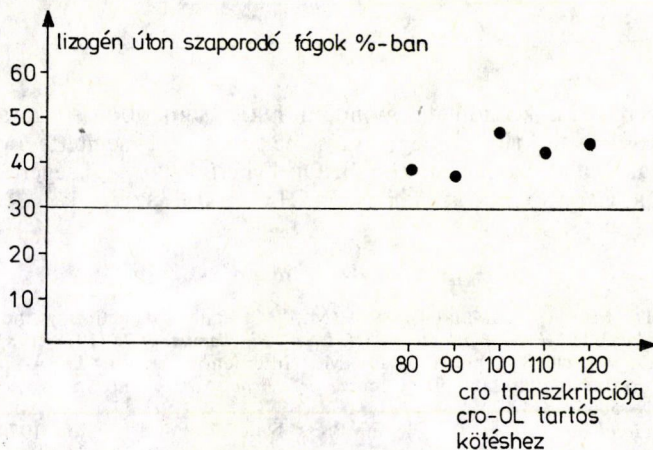
Ezután azt a paramétert változtattuk, hogy a Q génterméknek hány transzkripció során kell megszintetizálódnia, hogy a fág litikusan szaporodjon. Esetenként ismét 100 fágfertőzés szimulálását végeztük el. Az eredményeket a 9. ábra tartalmazza. Az ábráról látható, hogy a fág viselkedése igen érzékenyen reagál a vizsgált paraméter változtatására. Ha a Q génterméknek a litikus útra téréshez szükséges transzkripciónak a számát a modellben 10-ről 9-re csökkentjük, már 35%-ban fog a modellünk lizogén úton szaporodni, ami jó egyezést mutat a természetben megfigyelt 30%-kal.



9. ábra

Így tekintettel a lambda fág szabályozásának összetettségére, valamint az adatok becsült voltára, a futási eredmények alapján a lambda fág regulációs modellt elfogadhatjuk.

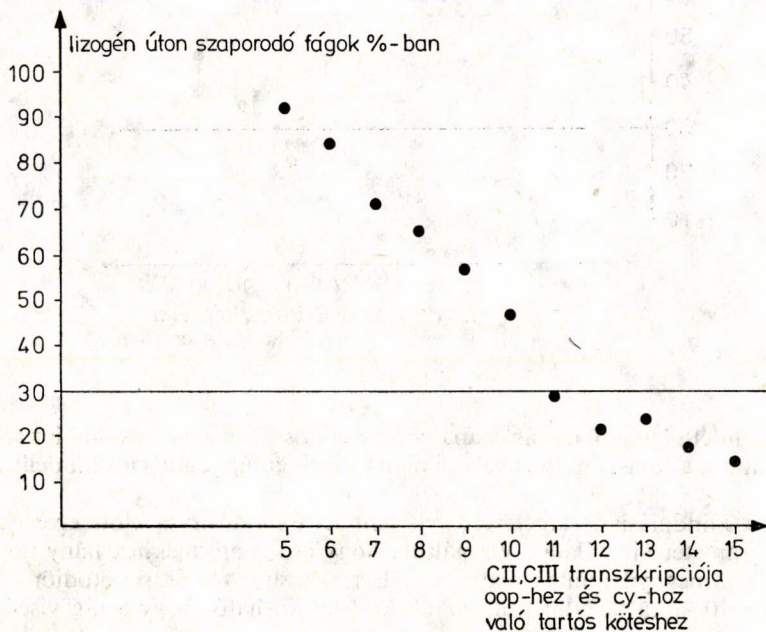
Még a kiindulási paraméterértékek fent leírt módosítása előtt elvégeztük két további paraméter vizsgálatát. Vizsgáltuk, hogy *cro* géntermékek hány transzkripció során kell megnyilvánulnia, hogy az OL szakaszhoz tartósan kötődjön. Az eredményeket a 10. ábra tartalmazza. Az ábráról leolvasható, hogy a fág viselkedése a vizsgált tartományban stabilitást mutat.



10. ábra

Vizsgáltuk továbbá, hogy a CII és CIII géntermékeknek hányszor kell megnyilvánulniuk, hogy az *oop* és *cy* szakaszokhoz tartósan kötődjenek. Az eredményeket a 11. ábra tartalmazza. Az ábráról látható, hogy a fág viselkedése érzékenyen reagál

a vizsgált paraméter változtatására. Megjegyezzük, hogy a modellen eszközölt változtatást ezen paraméter változtatásával is végre lehetett volna hajtani, amennyiben a CII és CIII gének szükséges megnyilvánulásait 10-ről 11-re növeljük.



11. ábra

Ezúton szeretnénk köszönetet mondani DR. CSIRIK JÁNOSnak, aki a kézirat többszöri átolvasásával, kritikai megjegyzéseinek tömegével segítette a dolgozat végleges formájának kialakulását. Köszönjük DR. MÁTÉ EÖRS segítségét. Köszönjük a konzultációt DR. CSÖRGŐ SÁNDORNak és DR. HATVANI LÁSZLÓnak.

IRODALOM

- [1] JOHNSON, A. D., MEYER, B. J. and PTASHNE, M., "Mechanism of action of the cro protein of bacteriophage lambda", *Proc. Natl. Acad. Sci. USA (Biochemistry)* **75** (1978) 1783—1787.
- [2] JOHNSON, A. D., MEYER, B. J. and PTASHNE, M., "Interactions between DNA-bound repressors govern regulation by the lambda phage repressor", *Proc. Natl. Acad. Sci. USA (Biochemistry)* **76** (1979) 5061—5065.
- [3] KING, R. C., *Handbook of Genetics* (Plenum Press, New York and London, 1974) Vol. I. 309—322.
- [4] KNUTH, D. E., *The Art of Computer Programming* (Addison-Wesley Publishing Company, 1969) Vol. II. 15—16, 35—40.
- [5] LEWIN, B., *Gene Expression-3 Plasmids and Phages* (John Wiley and Sons, 1977) 274—411.
- [6] OPPENHEIM, A. and OPPENHEIM, A., "Regulation of the int gene of bacteriophage lambda: Activation by the CII and CIII gene products and the role of the P_I and P_L promoters", *Molec. gen. Genet.* **165** (1978) 39—46.

- [7] RAY, U. and SKALKA, A., "Lysogenization of *E. coli* by phage lambda: Complementary activity of the host's DNA polymerase I and ligase and phage replication proteins O and P", *Journal of Virol.* **18** (1976) 511—517.

(Beérkezett: 1982. február 22.)

POLNER GÁBOR, DR. DALLMANN GÉZA ÉS DR. OROSZ LÁSZLÓ
JATE GENETIKAI TANSZÉK
6726 SZEGED, KÖZÉPFASOR 52.

MODELLING OF THE BACTERIOPHAGE LAMBDA REGULATION

G. POLNER, G. DALLMANN and L. OROSZ

Bacteriophage lambda is one of the most simple and definitely one of the best known living beings. Its multiplication may take place in two ways channelled by a complex regulation system. This paper describes the modelling of this regulation.

The analytical approach leads to a differential system of the type of *Kolmogorov backward*. This system can be solved theoretically but it is so complicated that in practice this way does not work.

Using the *Monte-Carlo computer simulation* we had the possibility to reconstruct the regulation of the phage from the function parameters measured by genetic experimentation. We studied how the system reacts upon altering a few function parameters. Unfortunately, much less experimental data are available to test our program than the number of possibilities to change the parameters, and that is why we expected that more than one constellation of the parameters could be produced which lead to an outcome compatible to the experimental data. Our results support these expectations.

POZITÍV DEFINIT KVADRATIKUS LJAPUNOV FÜGGVÉNYEK ALKALMAZÁSA STABILITÁSI VIZSGÁLATOKHOZ

KERTÉSZ VIKTOR

Budapest

Nemlineáris nem-autonóm differenciálegyenlet rendszerek megoldásainak normáira alsó és felső becsléseket adunk. Lineáris egyenleteknél a becslések speciális esete a W. COPPEL által közölt becslés [1].

A megadott becslések a változó alkalmas transzformációja esetén válnak megfelelően jókká. A jó transzformáló mátrix megkeresésének feladata lényegében ugyanaz, mint használható kvadratikusan *Ljapunov függvény* konstruálása.

Ha a nemlineáris rendszert olyan lineáris rendszer perturbációja révén létrejött rendszernek foghatjuk fel, amely lineáris rendszer együttható mátrixa saját értékei legnagyobb valós részének *improprius integrálja* nem végtelen, akkor bizonyos esetekben az azonosan zérus egyensúlyi helyzet stabilis. Ilyen esetekben a bemutatott becslések jól használhatók stabilitási kritériumok kimondásához, a *Ljapunov függvény* (transzformáló mátrix) az időben változó lineáris rendszerhez ugyanúgy konstruálható, mint autonóm lineáris rendszerhez.

Az eljárást a csillapított rezgőmozgás példáján mutatjuk be.

1. Differenciálegyenlet megoldások normájára vonatkozó alsó és felső becslések

Tekintsük az

$$(1.1) \quad \dot{y} = f(t, y); \quad y \in \mathbb{R}^n; \quad f: I \times D \mapsto \mathbb{R}^n;$$

$$I = [t_0; \infty); \quad D \subset \mathbb{R}^n, \quad D \text{ nyílt}$$

differenciálegyenlet rendszert. A továbbiakban ezt a rendszert az alábbi feltételek és jelölések mellett fogjuk vizsgálni:

$$(1.2) \quad f(t, y) = F(t, y)y = (A(t) + B(t, y))y,$$

$F(t, y)$, $A(t)$ és $B(t, y)$ folytonos $n \times n$ -es mátrix függvények $I \times T$ -ben, ahol

$$(1.3) \quad T = \{y \in \mathbb{R}^n: 0 \leq \varrho_1 < |y| < \varrho_2 \leq \infty\},$$

ϱ_1 ; ϱ_2 konstansok.

$y(t)$ (1.1)-nek azt az egyébként tetszőleges megoldását jelöli, amelyre $y(t) \in T$, ha $t_0 \leq t \leq t_1 \leq \infty$. Tételezzük fel, hogy ilyen megoldás létezik.

Az

$$(1.4) \quad \dot{x} = A(t)x; \quad t \geq t_0$$

egyenlet tetszőleges megoldása legyen $x(t)$, és alaprendszere $X(t)$.

Legyen $|\cdot|$ valamely vektornorma. Rendeljük ehhez a szokásos módon az

$$|A| = \sup_{|x|=1} |Ax|$$

mátrixnormát. Ezt a mátrixnormát felhasználva egy mátrixon értelmezett valós értékű függvény definícióját adjuk meg W. COPPEL [1] nyomán.

$$(1.5) \quad \mu(A) = \lim_{h \rightarrow 0+0} \frac{|I+Ah| - 1}{h}.$$

Az euklideszi vektornormából kiindulva ($|x| = (x_1^2 + \dots + x_n^2)^{1/2}$) és valós A mátrixok esetén $\mu(A)$ értéke nem más, mint A szimmetrikus részének $\left(\frac{1}{2}(A+A^T)\right)$ legnagyobb sajátértéke, amely a következőképpen is kifejezhető:

$$(1.6) \quad \mu(A) = \sup_{|x| \neq 0} \frac{1}{2} \frac{x^T(A+A^T)x}{x^T x},$$

ahol a felső „T” index a transzponáltat jelöli. Ugyanakkor $-\mu(-A)$ A szimmetrikus részének legkisebb sajátértékét adja:

$$(1.7) \quad -\mu(-A) = \inf_{|x| \neq 0} \frac{1}{2} \frac{x^T(A+A^T)x}{x^T x}.$$

Az euklideszi normával definiált μ függvény néhány egyéb tulajdonsága, amelyek (1.6)-ból és (1.7)-ből egyszerűen következnek:

$$\mu(A + \lambda I) = \mu(A) + \lambda,$$

$$\mu(A + B) \leq \mu(A) + \mu(B),$$

$$\mu(A) \leq |A|,$$

$$\mu(A^T) = \mu(A).$$

W. COPPEL tetszőleges vektornormával generált mátrixnorma felhasználásával definiált μ függvénnyel az (1.4) differenciálegyenlet $x(t)$ megoldásaira az alábbi becslést adja:

$$(1.8) \quad e^{\int_s^t \mu(-A(\tau)) d\tau} \leq \frac{|x(t)|}{|x(s)|} \leq e^{\int_s^t \mu(A(\tau)) d\tau}.$$

A továbbiakban e becslést javítjuk és nemlineáris esetre is kiterjesztjük. Becsléseinkhez az euklideszi vektor és mátrixnormát és az ehhez tartozó μ függvényt fogjuk használni.

Ezek után rátérhetünk a megoldások normájára vonatkozó alsó és felső becslés első tételére.

1.1. TÉTEL. Legyen $S(t)$ $[t_0, t_1]$ -en értelmezett tetszőleges reguláris és folytonosan differenciálható $n \times n$ -es mátrix függvény. Akkor (1.1) $y(t)$ megoldásaira (az (1.1)-re vonatkozó feltételek és jelölések mellett) az alábbi becslések érvényesek

$t_0 \leq s \leq t \leq t_1$ -ben:

$$(1.9) \quad |S^{-1}(s)|^{-1} |S(t)|^{-1} e^{-\int_s^t \mu(-S(\tau)F(\tau, y(\tau))S^{-1}(\tau) - \dot{S}(\tau)S^{-1}(\tau))d\tau} \leq \frac{|y(t)|}{|y(s)|},$$

$$(1.10) \quad \frac{|y(t)|}{|y(s)|} \leq |S^{-1}(t)| |S(s)| e^{\int_s^t \mu(S(\tau)F(\tau, y(\tau))S^{-1}(\tau) + \dot{S}(\tau)S^{-1}(\tau))d\tau}.$$

Bizonyítás. Tekintsük az

$$\dot{y}(t) = F(t, y(t))y(t); \quad t_0 \leq t \leq t_1$$

azonosságot. $|y(t)|^2$ deriváltjára kétféle kifejezést kaphatunk. Egyrészt:

$$\frac{d}{dt} |y(t)|^2 = 2|y(t)| \frac{d}{dt} |y(t)|,$$

másrészt:

$$\frac{d}{dt} |y(t)|^2 = \frac{d}{dt} y^T(t)y(t) = y^T(t) [F^T(t, y(t)) + F(t, y(t))] y(t),$$

ahol a felső „ T ” index — mint fenn — a transzponáltat jelöli. A két kifejezést egyenlővé téve és integrálva:

$$\frac{|y(t)|}{|y(s)|} = e^{\frac{1}{2} \int_s^t \frac{y^T(\tau) [F^T(\tau, y(\tau)) + F(\tau, y(\tau))] y(\tau)}{y^T(\tau)y(\tau)} d\tau}.$$

Alkalmazva a nyilvánvaló

$$-\mu(-F(\tau, y(\tau))) \leq \frac{y^T(\tau) [F^T(\tau, y(\tau)) + F(\tau, y(\tau))] y(\tau)}{2y^T(\tau)y(\tau)} \leq \mu(F(\tau, y(\tau)))$$

egyenlőtlenséget, kapható, hogy

$$(1.11) \quad e^{-\int_s^t \mu(-F(\tau, y(\tau)))d\tau} \leq \frac{|y(t)|}{|y(s)|} \leq e^{\int_s^t \mu(F(\tau, y(\tau)))d\tau}.$$

Ezek után $y(t)$ helyett vizsgáljuk a

$$z(t) = S(t)y(t)$$

függvényt, amelyre:

$$(1.12) \quad \dot{z}(t) = [S(t)F(t, y(t))S^{-1}(t) + \dot{S}(t)S^{-1}(t)]z(t) = G(t, y(t))z(t).$$

Más szóval: az y változót (1.1)-ben a következő módon transzformáljuk:

$$(1.13) \quad z = S(t)y.$$

Az (1.1)-re végrehajtott eljárást (amelynek eredménye (1.11) lett) alkalmazzuk most (1.12)-re. Ekkor a következőt kapjuk:

$$|z(s)| e^{-\int_s^t \mu(-G(\tau, y(\tau)))d\tau} \leq |z(t)| \leq |z(s)| e^{\int_s^t \mu(G(\tau, y(\tau)))d\tau}.$$

Figyelembe véve továbbá, hogy

$$|z(t)||S(t)|^{-1} \leq |y(t)| \leq |z(t)||S^{-1}(t)|,$$

$$|y(s)||S^{-1}(s)|^{-1} \leq |z(s)| \leq |y(s)||S(s)|,$$

a tétel állítása adódik!

Megjegyzés. Az (1.9) és (1.10) becslésekben (1.6) és (1.7) helyett (vagyis itt $\mu(G)$ és $-\mu(-G)$ helyett) (1.3) alapján célszerűbb a

$$\sup_{e_1 < |y| < e_2} \frac{y^T(G+G^T)y}{2y^Ty}; \quad \inf_{e_2 < |y| < e_1} \frac{y^T(G+G^T)y}{2y^Ty}$$

kifejezéseket használni. Az $y(t) \in T$, $t_0 \leq t \leq t_1$ relációt esetleg éppen (1.9) és (1.10)-zel bizonyíthatjuk (l. 1.2. példa, továbbá 2.1. és 2.2. tétel).

(1.9)-ben és (1.10)-ben $S(t)$ helyett alkalmazzuk $S_0(t) = v(t)S(t)$ -t, ahol $v(t)$ folytonosan deriválható pozitív skalár függvény.

Ekkor

$$S_0^{-1}(t) = \frac{1}{v(t)} S^{-1}(t),$$

$$\dot{S}_0(t) S_0^{-1}(t) = \frac{\dot{v}(t)}{v(t)} I + \dot{S}(t) S^{-1}(t),$$

$$|S_0(t)|^{-1} = \frac{1}{v(t)} |S(t)|^{-1},$$

$$|S_0^{-1}(s)|^{-1} = v(s) |S^{-1}(s)|^{-1}$$

Mindezekkel:

$$\begin{aligned} & |S^{-1}(s)|^{-1} |S(t)|^{-1} e^{-\int_s^t \mu(-S(\tau)F(\tau, y(\tau))S^{-1}(\tau) - \dot{S}(\tau)S^{-1}(\tau))d\tau} = \\ & = |S_0^{-1}(s)|^{-1} |S_0(t)|^{-1} e^{-\int_s^t \mu(-S_0(\tau)F(\tau, y(\tau))S_0^{-1}(\tau) - \dot{S}_0(\tau)S_0^{-1}(\tau))d\tau}; \\ & |S^{-1}(t)||S(s)| e^{\int_s^t \mu(S(\tau)F(\tau, y(\tau))S^{-1}(\tau) + \dot{S}(\tau)S^{-1}(\tau))d\tau} = \\ & = |S_0^{-1}(t)||S_0(s)| e^{\int_s^t \mu(S_0(\tau)F(\tau, y(\tau))S_0^{-1}(\tau) + \dot{S}_0(\tau)S_0^{-1}(\tau))d\tau}. \end{aligned}$$

Ebből következik, hogy ha $|S(t)|$, illetve $|S^{-1}(t)|$ is folytonosan deriválható ($S(t)$ folytonosan deriválhatóságából ez nem következik), akkor az $|S(t)|=1$, illetve $|S^{-1}(t)|=1$ feltételezés semmit nem ront az (1.9) és (1.10) becsléseken.

Az (1.9) és (1.10) becslésnek speciális esete az (1.11) becslés ($S(t)=I$), amely azonban általában túlságosan durva. Ez a becslés egyébként lineáris esetben ($B(t, y)=0$, $F(t, y)=A(t)$) éppen (1.8). Ha alkalmas $S(t)$ transzformáló mátrixot használunk, akkor (1.9) és (1.10) lényegesen jobb lehet (1.11)-nél.

1.1. PÉLDA. Legyen

$$F(t, y) = F = \begin{bmatrix} -1 & 0 \\ 8 & -7 \end{bmatrix} \quad \text{és} \quad S(t) = I.$$

F stabilis mátrix, tehát $\lim_{t \rightarrow \infty} |y(t)| = 0$, ugyanakkor $\mu(F) = 1$.

Ezzel (1.11)-ből az

$$|y(t)| \leq |y(s)| e^{t-s}$$

igen durva becslés adódik.

Legyen ezzel szemben $S(t) = S = \begin{bmatrix} 1 & 0 \\ -8 & 6 \end{bmatrix}$. Ezzel

$$SFS^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & -7 \end{bmatrix}; \quad \mu(SFS^{-1}) = -1; \quad |S||S^{-1}| \leq 17,$$

amellyel

$$|y(t)| \leq |y(s)| 17e^{-(t-s)}.$$

A fenti példa általánosítható. Konstans $F(t, y)$ esetében [2]-ből kiderül, hogy megadható olyan konstans S mátrix, amellyel az (1.9) és (1.10) becslés — bizonyos ott megfogalmazott értelemben — a lehető legélesebb.

1.2. PÉLDA. Legyen

$$(1.14) \quad F(t, y) = \begin{bmatrix} 0 & 1 \\ -1 & -p(t, y) \end{bmatrix},$$

ahol $p(t, y) \geq 0$ és folytonos. $S(t) = I$ választással az alábbi becslések adódnak:

$$(1.15) \quad e^{-\int_s^t p(\tau, y(\tau)) d\tau} \leq \frac{|y(t)|}{|y(s)|} \leq 1,$$

ugyanis

$$\mu \left(\begin{bmatrix} 0 & 1 \\ -1 & -p \end{bmatrix} \right) = 0,$$

$$-\mu \left(- \begin{bmatrix} 0 & 1 \\ -1 & -p \end{bmatrix} \right) = -p.$$

Ha

$$(1.16) \quad p(t, y) \leq p_0(t),$$

akkor

$$(1.17) \quad e^{-\int_s^t p_0(\tau) d\tau} \leq \frac{|y(t)|}{|y(s)|}.$$

Megállapítható, hogy $|y(t)|$ monoton csökkenő, tehát az $y(t) \equiv 0$ megoldás stabilis. Jó kritérium adható arra, hogy ez az egyensúlyi helyzet mikor nem aszimptotikusan stabilis: az $y(t) \equiv 0$ megoldás nem aszimptotikusan stabilis, ha létezik $p_0(t)$, hogy

$$(1.16) \text{ fennáll és } \int_s^\infty p_0(\tau) d\tau < \infty.$$

Az (1.11) becslés az aszimptotikus stabilitásra azonban itt semmiféle kritériumot nem ad. A későbbiekben látni fogjuk, hogy alkalmas $S(t)$ mátrixszal kritériumok adhatók az aszimptotikus stabilitásra is.

Adott $S(t)$ esetében vezessük be a

$$(1.18) \quad v(t, y) = y^T S^T(t) S(t) y = y^T V(t) y$$

Ljapunov függvényt, amelynek az (1.1) rendszer szerinti deriváltja:

$$(1.19)$$

$$\dot{v}_{(1.1)}(t, y) = \frac{d}{dt} v(t, y(t)) = y^T(t) [V(t) F(t, y(t)) + F^T(t, y(t)) V(t) + \dot{V}(t)] y(t),$$

ahol $y = y(t)$.

Egyszerűen kimutathatók az alábbi összefüggések:

$$(1.20) \quad 2\mu(S(t)F(t, y(t))S^{-1}(t) + \dot{S}(t)S^{-1}(t)) = \sup_{|y| \neq 0} \frac{\dot{v}_{(1.1)}(t, y)}{v(t, y)},$$

$$(1.21) \quad -2\mu(-S(t)F(t, y(t))S^{-1}(t) - \dot{S}(t)S^{-1}(t)) = \inf_{|y| \neq 0} \frac{\dot{v}_{(1.1)}(t, y)}{v(t, y)}.$$

Ezekkel (1.9) és (1.10)-nek megfelelő becslések:

$$(1.22) \quad e^{\int_s^t \inf_{|y| \neq 0} \frac{\dot{v}_{(1.1)}(\tau, y)}{v(\tau, y)} d\tau} \leq \frac{v(t, y(t))}{v(s, y(s))} \leq e^{\int_s^t \sup_{|y| \neq 0} \frac{\dot{v}_{(1.1)}(\tau, y)}{v(\tau, y)} d\tau}.$$

Másrésről, ha adott $V(t)$ $[t_0, t_1]$ -ben értelmezett pozitív definit és folytonosan deriválható mátrix, akkor felírható $V(t) = S^T(t) S(t)$ alakban, ahol $S(t)$ az 1.1 tétel szerinti.

A fenti összefüggések alapján indokolt, hogy ne tegyünk különbséget jó becslésre alkalmas transzformáló mátrix, illetőleg kvadrátikus *Ljapunov függvény* konstruálásának feladata között.

A továbbiakban még egy alsó és felső becslést adunk a megoldások normájára. Ehhez a következő lemmára lesz szükségünk.

1.1. LEMMA. Ha valamely rögzített $s \geq t_0$ -ra létezik $0 < a(t)$ függvény, hogy

$$(1.23) \quad |x(t)| \leq |x(s)| a(t); \quad t \geq s,$$

akkor

$$(1.24) \quad |X(t)| \leq |X(s)| a(t),$$

ahol $x(t)$ és $X(t)$ (1.4) tetszőleges megoldása, illetve alaprendszer.

Bizonyítás. $x(t)$ ilyen alakban írható:

$$x(t) = X(t) X^{-1}(s) x(s).$$

Továbbá

$$\frac{|x(t)|}{|x(s)|} = \frac{|X(t) X^{-1}(s) x(s)|}{|x(s)|} \leq a(t)$$

igaz bármely $x(s)$ -re ($|x(s)| \neq 0$), ezzel:

$$(1.25) \quad |X(t)X^{-1}(s)| \leq a(t).$$

(1.24) annak figyelembevételével adódik, hogy

$$|X(t)| = |X(t)X^{-1}(s)X(s)| \leq |X(t)X^{-1}(s)||X(s)|.$$

Az

$$(1.26) \quad \dot{u} = -A^T(t)u$$

rendszer alaprendszere $X^{T-1}(t)$, ha $X(t)$ (1.4) alaprendszere. Felhasználva az 1.1. lemmát és (1.4)-re, valamint (1.26)-ra alkalmazva (1.8)-at:

$$(1.27) \quad |X(t)|^{-1} \leq |X(s)|^{-1} e^{-\int_s^t \mu(A(\tau)) d\tau},$$

$$(1.28) \quad |X^{-1}(t)| \leq |X^{-1}(s)| e^{\int_s^t \mu(-A(\tau)) d\tau}.$$

Legyen $Q(t) = \dot{S}(t)S^{-1}(t)$, akkor most $S(t)$ az

$$(1.29) \quad \dot{s} = Q(t)s$$

differenciálegyenlet alaprendszere. A fenti eljárást ((1.27) és (1.28)) erre az esetre alkalmazva:

$$(1.30) \quad |S(t)|^{-1} \leq |S(s)|^{-1} e^{-\int_s^t \mu(Q(\tau)) d\tau},$$

$$(1.31) \quad |S^{-1}(t)| \leq |S^{-1}(s)| e^{\int_s^t \mu(-Q(\tau)) d\tau}.$$

Ezzel (1.9) és (1.10) helyett a következő írható:

$$(1.32) \quad |S^{-1}(s)|^{-1} |S(s)|^{-1} e^{\int_s^t [-\mu(-S(\tau)F(\tau, y(\tau))S^{-1}(\tau) - Q(\tau)) - \mu(Q(\tau))] d\tau} \leq \frac{|y(t)|}{|y(s)|},$$

$$(1.33) \quad \frac{|y(t)|}{|y(s)|} \leq |S^{-1}(s)| |S(s)| e^{\int_s^t [\mu(S(\tau)F(\tau, y(\tau))S^{-1}(\tau) + Q(\tau)) + \mu(-Q(\tau))] d\tau}.$$

Ismeretes, hogy az (1.4) egyenletnél $A(t)$ sajátértékei valós részének negatív volta még nem garantálja az $x \equiv 0$ megoldás aszimptotikus stabilitását vagy akár stabilitását.

Tegyük fel, hogy $A(t)$ rögzített t -re stabilis és $S(t) A(t) S^{-1}(t) A(t)$ -nek ε -normál formája [2]. (1.33)-ból megállapítható, hogy — eltekintve a $B(t, y)$ nem lineáris perturbációtól — elsősorban nem a sajátértékek idő függése az oka annak, ha az $x \equiv 0$ megoldás nem stabilis, hanem a sajátvektorok nem konstans volta. Vagyis:

$$\mu(Q) + \mu(-Q) > 0.$$

(Vegyük még figyelembe, hogy ha $S(t)$ állandó lenne, akkor $Q(t) \equiv 0$ és $\mu(Q) = \mu(-Q) = 0$ lenne.)

1.3. PÉLDA. Ha $A(t)$ szimmetrikus, akkor $\mu(A(t))$ $A(t)$ -nek a legnagyobb saját értékét adja, amely legyen $\lambda(t)$, $-\mu(-A(t))$ pedig a $v(t)$ legkisebb saját értéket. Legyen $S=I$. Ekkor $Q=0$.

Vegyük figyelembe, hogy

$$\begin{aligned}\mu(A(t) + B(t, y)) &\leq \mu(A(t)) + |B(t, y)|, \\ -\mu(-A(t) - B(t, y)) &\leq -\mu(-A(t)) - |B(t, y)|.\end{aligned}$$

Legyen $t_1 = \infty$, $q_1 = 0$ és

$$|B(t, y)| \leq \delta(t).$$

Az $x \equiv 0$ megoldás aszimptotikusan stabilis, ha $\lim_{t \rightarrow \infty} \int_s^t (\lambda(\tau) + \delta(\tau)) d\tau = -\infty$. Az $x \equiv 0$ megoldás stabilis, ha

$$\lim_{t \rightarrow \infty} \int_s^t (\lambda(\tau) + \delta(\tau)) d\tau < \infty.$$

Végül $x \equiv 0$ nem aszimptotikusan stabilis ha,

$$\lim_{t \rightarrow \infty} \int_s^t (v(\tau) - \delta(\tau)) d\tau > \infty.$$

1.4. PÉLDA. Legyen A_0 konstans stabilis mátrix, de nem szimmetrikus és $\mu(A_0) > 0$. Legyen Z konstans antiszimmetrikus, hogy $A_0 + Z$ -nek már legyen pozitív valós részű saját értéke. Z -hez mindig van $M(t)$ ortogonális mátrix ($M^{-1}(t) = M^T(t)$), hogy $\dot{M}(t) M^T(t) = Z$. $M(t)$ a $\dot{z} = Zz$ differenciálegyenlet egyik alaprendszer. Tekintsük az

$$\dot{y} = (A_0 + Z)y$$

differenciálegyenletet és az

$$x = M^T y$$

transzformációt. x -re az alábbi egyenlet adódik:

$$\dot{x} = M^T A_0 M x = A(t)x.$$

Ekkor $A(t)$ saját értékei konstansok, valós részük negatív, ugyanakkor van olyan megoldás, hogy $\lim_{t \rightarrow \infty} |x(t)| = \infty$. Legyen $S_0 A_0 S_0^{-1}$ Jordan-féle normál forma, ahol S_0 konstans. Akkor most x -re az $S(t) = S_0 M(t)$ transzformációt alkalmazva ($z = S(t)x$) $Q = \dot{S}(t)S^{-1}(t) = S_0 \dot{M}(t)M^T(t)S_0^{-1} = S_0 Z S_0^{-1}$ nyilván olyan lesz, hogy

$$\mu(S(t)A(t)S^{-1}(t)) < 0,$$

azonban

$$\mu(-Q) + \mu(Q) > 0,$$

sőt

$$\mu(S(t)A(t)S^{-1}(t)) + \mu(-Q) + \mu(Q) > 0.$$

Legyen $A(t)$ legnagyobb valós részű saját értékének valós része $\alpha(t)$. A csillapított rezgőmozgás meglepő és sokat ígérő példát szolgáltat arra, hogy $\lim_{t \rightarrow \infty} \int_s^t \alpha(\tau) d\tau = -\infty$ esetében az $x \equiv 0$ megoldás aszimptotikusan stabilis, sőt az aszimptotikus stabilitást mennyire nagy perturbációk sem szüntetik meg.

2. A csillapított rezgőmozgás stabilitási vizsgálata

Az 1. pont (1.10) becslésének önmagában is fontos alkalmazási példaként tárgyaljuk a csillapított rezgőmozgás egy differenciálegyenletét.

Tekintsük az

$$(2.1) \quad u'' + p(t, u, u')u' + u = 0; \quad p(t, u, u') \cong 0$$

($p(t, u, u')$ folytonos) differenciálegyenletet, amelyet

$$(2.2) \quad y = \begin{bmatrix} u \\ u' \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

bevezetésével (1.1) alakra hozhatunk, ahol $F(t, y)$ (1.14) szerinti. Más jelöléssel

$$(2.3) \quad F(p) = \begin{bmatrix} 0 & 1 \\ -1 & -p(t, y) \end{bmatrix}.$$

Legyen $0 \leq \delta(t) < 2$ folytonosan deriválható függvény és

$$(2.4) \quad F(\delta) = \begin{bmatrix} 0 & 1 \\ -1 & -\delta \end{bmatrix}.$$

$S(\delta)$ legyen olyan mátrix, amely $F(\delta)$ -t valós *Jordan-normál* alakra hozza:

$$(2.5) \quad S(\delta)F(\delta)S^{-1}(\delta) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

ahol $a + ib$, ill. $a - ib$ $F(\delta)$ saját értékei. Jelen esetben $a = -\delta/2$.

$$(2.6) \quad S(\delta) = \begin{bmatrix} 0 & \alpha \\ \frac{2}{\beta} & \frac{\delta}{\beta} \end{bmatrix},$$

ahol

$$(2.7) \quad \alpha(t) = \sqrt{2 + \delta(t)} \\ \beta(t) = \sqrt{2 - \delta(t)},$$

$S(\delta)F(p)S^{-1}(\delta) + \dot{S}(\delta)S^{-1}(\delta)$ szimmetrikus része:

$$(2.8) \quad \frac{S(\delta)F(p)S^{-1}(\delta) + \dot{S}(\delta)S^{-1}(\delta) + [S(\delta)F(p)S^{-1}(\delta) + \dot{S}(\delta)S^{-1}(\delta)]^T}{2} = \\ = \begin{bmatrix} \frac{\dot{\delta}}{2\alpha^2} - \frac{\delta}{2} + (\delta - p) & \frac{\dot{\delta}}{2\omega} + \frac{\delta}{2\omega}(\delta - p) \\ \frac{\dot{\delta}}{2\omega} + \frac{\delta}{2\omega}(\delta - p) & \frac{\dot{\delta}}{2\beta^2} - \frac{\delta}{2} \end{bmatrix},$$

ahol

$$(2.9) \quad \omega = \alpha\beta.$$

A (2.8) mátrix legnagyobb saját értéke:

$$(2.10) \quad \mu(S(\delta)F(p)S^{-1}(\delta) + \dot{S}(\delta)S^{-1}(\delta)) = -\frac{p}{2} + \frac{\delta}{\omega^2} + \sqrt{\frac{(\delta-p)^2}{\omega^2} + \frac{\dot{\delta}^2}{\omega^4}}.$$

Az utóbbi kifejezést jelöljük egyszerűen $\mu(\delta, \dot{\delta}, p)$ -vel. Nyilván:

$$(2.11) \quad \mu(\delta, \dot{\delta}, p) \leq -\frac{p}{2} + \frac{|\delta-p|}{\omega} + \frac{\dot{\delta} + |\dot{\delta}|}{\omega^2}.$$

Vezessük be továbbá a következő jelöléseket:

$$(2.12) \quad \mu_0(\delta, p) = -\frac{p}{2} + \frac{|\delta-p|}{\omega},$$

$$(2.13) \quad \Delta\mu(\delta, \dot{\delta}) = \frac{\dot{\delta} + |\dot{\delta}|}{\omega^2} = \begin{cases} 0, & \text{ha } \dot{\delta} \leq 0 \\ \frac{2\dot{\delta}}{\omega^2}, & \text{ha } \dot{\delta} \geq 0. \end{cases}$$

Tekintettel arra, hogy

$$\frac{d}{dt} \ln \frac{\alpha}{\beta} = \frac{2\dot{\delta}}{\omega^2},$$

ezért

$$(2.14) \quad \int_{t_a}^{t_b} \Delta\mu(\delta(\tau), \dot{\delta}(\tau)) d\tau = \frac{\alpha(t_b)}{\beta(t_b)} \frac{\beta(t_a)}{\alpha(t_a)},$$

ahol $[t_a, t_b]$ -ben $\dot{\delta} \geq 0$.

$S(\delta)$ és $S^{-1}(\delta)$ normájára az alábbi kifejezés kapható:

$$|S(\delta)| = \sqrt{2} \frac{\alpha}{\beta},$$

$$|S^{-1}(\delta)| = \frac{1}{\sqrt{2}},$$

amelyekből:

$$(2.15) \quad |S(\delta(t_0))| |S^{-1}(\delta(t))| = \frac{\alpha(t_0)}{\beta(t_0)}.$$

$\delta(t)$ $[t_0, t_1]$ -ben a következő tulajdonságok valamelyikével rendelkezzen:

- (a) $\delta(t)$ monoton csökkenő
- (b) $\delta(t)$ monoton növekedő
- (c) létezik $t_0 = t_{m0} \leq t_{M1} < t_{m1} < t_{M2} < t_{m2} < \dots < t_{Mn} < t_{mn} < \dots$,
hogy $[t_{Mi}; t_{mi}]$ -ben $\delta(t)$ monoton csökkenő és
 $[t_{mi}; t_{Mi+1}]$ -ben $\delta(t)$ monoton növekedő.

Akkor (2.11); (2.12); (2.13); (2.14) és (2.15) figyelembevételével az (1.10) felső becslés eképpen alakul:

$$(2.16) \quad (a) \quad \frac{|y(t)|}{|y(t_0)|} \leq \frac{\alpha(t_0)}{\beta(t_0)} e^{\int_{t_0}^t \mu_0(\delta(\tau), p(\tau, y(\tau))) d\tau},$$

$$(b) \quad \frac{|y(t)|}{|y(t_0)|} \leq \frac{\alpha(t)}{\beta(t)} e^{\int_{t_0}^t \mu_0(\delta(\tau), p(\tau, y(\tau))) d\tau},$$

$$(2.17) \quad (c) \quad \frac{|y(t)|}{|y(t_0)|} \leq \frac{\alpha(t_{Mn})}{\beta(t_{Mn})} \prod_{i=1}^{n-1} \frac{\alpha(t_{Mi})}{\beta(t_{Mi})} \frac{\beta(t_{mi})}{\alpha(t_{mi})} e^{\int_{t_0}^t \mu_0(\delta(\tau), p(\tau, y(\tau))) d\tau},$$

ahol $t_{Mn} \leq t \leq t_{mn} \leq t_1$.

Természetesen $y(t)$ nem ismeretes, így a fenti becslések csak akkor válnak használhatóvá, ha $\mu_0(\delta(t), p(t, y(t)))$ -re felső becslést adunk. Eléggő durva, de igen jól kezelhető becslés az alábbi.

Válasszuk $\delta(t)$ -t úgy (ha ez egyáltalán lehetséges), hogy $|y| \leq |y(t_0)|$ -re és $[t_0, t_1]$ -ben

$$(2.18) \quad 0 \leq p(t, y(t)) \leq \frac{1}{\delta(t)},$$

$$(2.19) \quad 0 \leq \delta(t) \leq 1.$$

Ekkor

$$(2.20) \quad \mu_0(\delta(t), p(t, y(t))) \leq c(t) \delta(t),$$

ahol

$$(2.21) \quad c(t) = \begin{cases} -\left(1 - \frac{1}{\sqrt{3}}\right), & \text{ha } \delta(t) \leq p(t, y) < \frac{1}{\delta(t)} \\ \frac{1}{\sqrt{3}}, & \text{ha } 0 \leq p(t, y) \leq \delta(t). \end{cases}$$

Az (2.16)–(2.21) összefüggések felhasználásával stabilitási kritériumok adhatók. Példaképpen két tételt közlünk.

2.1. TÉTEL. A (2.1) differenciálegyenlet $u=v'=0$ egyensúlyi helyzete aszimptotikusan stabilis ha létezik olyan $\delta(t)$ valós függvény, amelyre $t \geq t_0$ -ban

- (i) $0 < \delta(t) \leq 1$;
- (ii) $\delta(t)$ folytonosan deriválható;
- (iii) $\dot{\delta}(t) \leq 0$;
- (iv) az origó $((0; 0))$ egy környezetében minden onnan induló megoldás létezik $[t_0, \infty)$ -ben és itt

$$0 \leq p(t, u, v) \leq \frac{1}{\delta(t)};$$

$$(v) \quad \lim_{t \rightarrow \infty} \int_{t_0}^t c_0(\tau) \delta(\tau) d\tau = \infty,$$

ahol

$$c_0(t) = \begin{cases} 1, & \text{ha } \delta(t) \leq p(t, u, v) \\ -\frac{1}{\sqrt{3}-1}, & \text{ha } p(t, u, v) < \delta(t). \end{cases}$$

2.2. KÖVETKEZMÉNY. A (2.1) differenciálegyenlet $u=v=0$ egyensúlyi helyzete aszimptotikusan stabilis, ha létezik olyan $\delta(t)$ valós függvény, amelyre $t \geq t_0$ -ban

- (i) $0 < \delta(t) \leq 1$;
- (ii) $\delta(t)$ folytonosan deriválható;
- (iii) $\dot{\delta}(t) \leq 0$;
- (iv) Az origó $((0; 0))$ egy környezetében minden onnan induló megoldás létezik $[t_0, \infty)$ -ben és itt

$$\delta(t) \leq p(t, u, v) \leq \frac{1}{\delta(t)}.$$

$$(v) \quad \lim_{t \rightarrow \infty} \int_{t_0}^t \delta(\tau) d\tau = \infty.$$

IRODALOM

- [1] COPPEL, W., *Stability and Asymptotic Behaviour of Differential Equation* (D. C. Heath, Boston, 1965).
- [2] GARAY, B. and KERTÉSZ, V., "Lozinsky's functional and the transformation of variables", *Zeitschrift für Analysis und ihre Anwendungen* 2 (1984). Megjelenés alatt.

(Beérkezett: 1983. február 25.)

KERTÉSZ VIKTOR

BME GÉPÉSZMÉRNÖKI KAR MATEMATIKA TANSZÉK
1521 BUDAPEST, STOCZEK U. H. ÉP. IV. E. 14.

APPLICATION OF POSITIVE DEFINITE QUADRATIC LYAPUNOV FUNCTIONS TO STABILITY INVESTIGATIONS

V. KERTÉSZ

We give lower and upper bounds for the norm of the solutions of nonlinear nonautonomous differential equations. In the case of linear equations these bounds are special cases of those given by W. COPPEL [1]. The bounds become appropriate by applying a variable transformation.

We show the procedure on the example of damped oscillation.

MEGJEGYZÉS AZ EGYENLŐTLENSÉG FELTÉTELEK MELLETTI MINIMALIZÁLÁSRÓL

RAPCSÁK TAMÁS

Budapest

A dolgozatban megmutatjuk, hogy az egyenlőtlenség feltételekkel korlátozott nemlineáris programozási feladat első és másodrendű, szükséges és elegendő optimalitási feltételei az egyváltozós függvények szélsőértékhelyein teljesülő első és másodrendű kritériumok természetes általánosításai.

1. Bevezetés

A [11] dolgozatban a

$$(1.1) \quad \min f(\mathbf{x}); \quad \mathbf{x} = \mathbf{x}(\mathbf{u}) \in R^n, \quad \mathbf{u} \in U_k; \quad f(\mathbf{x}), \quad x_i(\mathbf{u}) \in C^2, \quad i = 1, \dots, n$$

feladattal foglalkoztunk, ahol az $\mathbf{x}(\mathbf{u})$ elemi felület *Gauss-féle paraméteres alakban* volt adva és U_k a k -dimenziós paraméter teret vagy annak egy részalmazát jelentette. Az $f(\mathbf{x})$ függvény egy az $\mathbf{x}(\mathbf{u})$ felületet tartalmazó nyílt halmazon volt értelmezve. Megmutattuk, hogy az egyváltozós függvények szélsőérték helyeire vonatkozó kritériumok természetes módon kiterjeszthetők erre a feladatra, s így megkaptuk az optimalitás első és másodrendű, szükséges és elegendő feltételeit. Ezután a

$$(1.2) \quad \min f(\mathbf{x}); \quad g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m; \quad \mathbf{x} \in R^n, \quad f(\mathbf{x}), \quad g_i(\mathbf{x}) \in C^2, \\ i = 1, \dots, m$$

feladatot tekintettük, ahol \mathbf{x}_0 volt a vizsgált pont, és feltételeztük, hogy a hiperfelületek metszete az \mathbf{x}_0 pont egy környezetében elemi felületet határoz meg és $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$. Ennek a felületnek az érintősíkját \hat{T} jelölte és bevezettük az alábbi jelölést is.

$$(1.3) \quad \hat{M} = \{\mathbf{y} | \nabla g_i(\mathbf{x}_0)\mathbf{y} = 0, \quad i = 1, \dots, m\}.$$

Azt mondjuk, hogy az (1.2) feladat \mathbf{x}_0 pontjában a *Lagrange feltétel* teljesül, ha léteznek $\mu_i, i = 1, \dots, m$ értékek úgy, hogy

$$(1.4) \quad \nabla f(\mathbf{x}_0) - \sum_{i=1}^m \mu_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}.$$

Az alábbi tételeket bizonyítottuk be.

1.1. TÉTEL. Annak szükséges és elegendő feltétele, hogy minden $f(\mathbf{x})$ függvény esetén, amelyre \mathbf{x}_0 az (1.2) feladat optimum pontja az (1.4) feltételek teljesüljenek az, hogy $\hat{M} = \hat{T}$ legyen.

Ebből a tételből következik, hogy az (1.2) feladat esetén az $\hat{M} = \hat{T}$ regularitási feltételnél nem lehet enyhébbet találni.

1.2. TÉTEL. Ha \mathbf{x}_0 az (1.2) feladat optimum pontja és $\hat{M} = \hat{T}$ akkor (1.4) teljesül és

$$(1.5) \quad \mathbf{y}^T \nabla^2 f(\mathbf{x}_0) \mathbf{y} - \sum_{i=1}^m \mu_i \mathbf{y}^T \nabla^2 g_i(\mathbf{x}_0) \mathbf{y} \equiv 0, \quad \mathbf{y} \in \hat{M}.$$

1.3. TÉTEL. Ha az (1.2) feladat \mathbf{x}_0 pontjában teljesül az (1.4) feltétel és

$$(1.6) \quad \mathbf{y}^T \nabla^2 f(\mathbf{x}_0) \mathbf{y} - \sum_{i=1}^m \mu_i \mathbf{y}^T \nabla^2 g_i(\mathbf{x}_0) \mathbf{y} > 0, \quad \mathbf{y} \in \hat{M},$$

akkor az \mathbf{x}_0 pont szigorú lokális minimum.

Az 1.2., 1.3. tételek az optimalitás jól ismert szükséges és elegendő feltételei [1], [7]. Azonban a tételeket nem a szokásos úton nyertük, hanem az egyváltozós függvényekre teljesülő optimalitási kritériumokat általánosítottuk a klasszikus differenciálgeometria eszközeinek a felhasználásával.

Ebben a dolgozatban azt mutatjuk meg, hogy az egyenlőtlenség feltételekkel megadott nemlineáris programozási feladat optimalitási kritériumai az itt tárgyalt esetben direkt módon származtathatók az 1.2., 1.3. tételekből, tehát a szükséges rész bizonyításánál nem használjuk fel a *Farkas tételt* [3], [4], az elegendőségi feltételek bizonyításánál pedig elkerüljük az indirekt bizonyítási technikát [1], [3], [7]. Így megmutatjuk azt is, hogy az egyenlőtlenség feltételekkel megadott nemlineáris programozási feladat optimalitási feltételei is az egyváltozós függvények minimumpontjaiban teljesülő első és másodrendű kritériumok természetes általánosításai.

2. A szükséges feltételek származtatása

A következő feladattal foglalkozunk:

$$(2.1) \quad \min f(\mathbf{x}); \quad g_i(\mathbf{x}) \equiv 0, \quad i = 1, \dots, m; \\ \mathbf{x} \in R^n, \quad f(\mathbf{x}), \quad g_i(\mathbf{x}) \in C^2, \quad i = 1, \dots, m.$$

Legyen \mathbf{x}_0 a vizsgált pont és $I(\mathbf{x}_0)$ az aktív indexek halmaza. Tekintsük most a (2.1) probléma helyett a

$$(2.2) \quad \min f(\mathbf{x}); \quad h_i(\mathbf{x}, \mathbf{z}) = g_i(\mathbf{x}) - \frac{1}{2} z_i^2 = 0, \quad i \in I(\mathbf{x}_0)$$

feladatot.

Mivel \mathbf{x}_0 akkor és csak akkor lokális optimuma a (2.1) problémának, ha $(\mathbf{x}_0, \mathbf{0})$ is lokális optimuma (2.2)-nek, ezért (2.1) helyettesíthető ez utóbbi feladattal.

Tegyük fel, hogy az $(\mathbf{x}_0, \mathbf{0})$ pont egy környezetében a $h_i(\mathbf{x}, \mathbf{z}) = 0$, $i \in I(\mathbf{x}_0)$ nívófelületek metszete elemi felületet határoz meg, az $I(\mathbf{x}_0)$ halmazban szereplő indexek száma k_1 és az elemi felület érintősíkját az $(\mathbf{x}_0, \mathbf{0})$ pontban T jelöli, valamint legyen

$$(2.3) \quad M = \{(\mathbf{y}_1, \mathbf{y}_2) / \nabla h_i(\mathbf{x}_0, \mathbf{0}) \cdot (\mathbf{y}_1, \mathbf{y}_2) = 0, \quad i \in I(\mathbf{x}_0)\}.$$

A (2.3) egyenlőségben szereplő $\mathbf{y}_1, \mathbf{y}_2$ vektorok n , illetve k_1 dimenziósak. Mivel

$\nabla h_i(\mathbf{x}_0, 0) = (\nabla g_i(\mathbf{x}_0), 0)$, $i \in I(\mathbf{x}_0)$, ezért

$$(2.4) \quad M = \{(\mathbf{y}_1, \mathbf{y}_2) / \nabla g_i(\mathbf{x}_0) \mathbf{y}_1 = 0, \quad i \in I(\mathbf{x}_0), \quad \mathbf{y}_2 \text{ tetszőleges}\}.$$

Alkalmazva az 1.2. tételt, azt kapjuk, hogy ha $(\mathbf{x}_0, 0)$ a (2.2) feladat optimum pontja és $M = T$, akkor léteznek μ_i , $i \in I(\mathbf{x}_0)$ értékek úgy, hogy

$$(2.5) \quad \nabla f(\mathbf{x}_0) - \sum_{i \in I(\mathbf{x}_0)} \mu_i \nabla g_i(\mathbf{x}_0) = 0.$$

Ha ehhez hozzávesszük a

$$(2.6) \quad \mu_i = 0, \quad i \notin I(\mathbf{x}_0)$$

egyenlőségeket, akkor a *Karush—Kuhn—Tucker-tétel* bizonyításához csak a μ_i , $i \in I(\mathbf{x}_0)$ szorzók nemnegativitását kell belátni.

Felhasználva az 1.2. tételben szereplő (1.5) másodrendű szükséges feltételt, kapjuk, hogy a

$$(2.7) \quad \nabla_{(\mathbf{x}, \mathbf{z})}^2 L(\mathbf{x}_0, 0, \boldsymbol{\mu}) = \begin{pmatrix} \nabla^2 f(\mathbf{x}_0) - \sum_{i \in I(\mathbf{x}_0)} \mu_i \nabla^2 g_i(\mathbf{x}_0) & & & \\ & \mu_1 & & \\ & & \mu_2 & \\ & & & \ddots & \\ & & & & \mu_i & \\ & & & & & \ddots & \\ & & & & & & \mu_{k_1} \end{pmatrix}$$

mátrix pozitív szemidefinit az M halmazon, ahol az $L(\mathbf{x}, \mathbf{z}, \boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i \in I(\mathbf{x}_0)} \mu_i h_i(\mathbf{x}, \mathbf{z})$ függvény a (2.2) probléma *Lagrange függvénye*. Ha most azokat az M halmazhoz tartozó vektorokat választjuk ki, amelyekre $\mathbf{y}_1 = \mathbf{0}$, $\mathbf{y}_2 = \mathbf{0}$ kivéve egy y_{2i} komponenst, $i \in I(\mathbf{x}_0)$, akkor éppen a *Lagrange szorzók* nemnegativitását kapjuk.

Ha $\mathbf{y}_2 = \mathbf{0}$, akkor az következik, hogy a $\nabla^2 f(\mathbf{x}_0) - \sum_{i \in I(\mathbf{x}_0)} \mu_i \nabla^2 g_i(\mathbf{x}_0)$ mátrix pozitív szemidefinit a $\nabla g_i(\mathbf{x}_0) \mathbf{y}_1 = 0$, $i \in I(\mathbf{x}_0)$ altéren.

Ezzel bebizonyítottuk az alábbi tételt.

2.1. TÉTEL. Ha \mathbf{x}_0 a (2.1) feladat lokális optimum pontja és $M = T$, akkor léteznek $\mu_i \geq 0$, $i = 1, \dots, m$ értékek úgy, hogy a

$$(2.8) \quad \nabla f(\mathbf{x}_0) - \sum_{i=1}^m \mu_i \nabla g_i(\mathbf{x}_0) = 0$$

$$\mu_i g_i(\mathbf{x}_0) = 0, \quad i = 1, \dots, m$$

$$\mathbf{y}^T \Delta^2(\mathbf{x}_0) \mathbf{y} - \sum_{i=1}^m \mu_i \mathbf{y}^T \nabla^2 g_i(\mathbf{x}_0) \mathbf{y} \geq 0, \quad \text{ha} \quad \nabla g_i(\mathbf{x}_0) \mathbf{y} = 0, \quad i \in I(\mathbf{x}_0)$$

feltételek teljesülnek.

Ezek a (2.1) feladatnál az optimalitás jól ismert szükséges feltételei [1].

Megjegyezzük, hogy a (2.2) feladatot tekintve a tételben szereplő regularitási feltételnél az 1.1. tétel szerint nem lehet enyhébbet találni.

Ismert az az állítás, hogy ha az M halmaz definíciójában szereplő gradienssek lineárisan függetlenek, akkor $M = T$ [7].

A következő részben az $M = T$ regularitási feltételt vizsgáljuk a (2.2) feladat esetén.

Az általánosság megszorítása nélkül feltehetjük, hogy a $\nabla g_i(\mathbf{x}_0)$, $i \in I(\mathbf{x}_0)$ vektorok között k darab független van. Az egyszerűség kedvéért legyenek ezek $\nabla g_i(\mathbf{x}_0)$, $i=1, \dots, k$. Azt mondjuk, hogy egy $h_i(\mathbf{x}, \mathbf{z})=0$, $i \in I(\mathbf{x}_0)$ feltétel az $(\mathbf{x}_0, \mathbf{0})$ pontban redundáns egy F halmazra nézve, ha $h_i(\mathbf{x}_0, \mathbf{0})=0$ és

$$(2.9) \quad h_i(\mathbf{x}, \mathbf{z}) = 0, \quad (\mathbf{x}, \mathbf{z}) \in F$$

esetén.

2.2. TÉTEL. Az $M=T$ regularitási feltétel akkor és csak akkor teljesül a (2.2) feladatban, ahol az $(\mathbf{x}_0, \mathbf{0})$ pont egy környezetében a $h_i(\mathbf{x}, \mathbf{z})=0$, $i \in I(\mathbf{x}_0)$ nivófelületek metszete elemi felületet határoz meg, ha van a feltételeknek olyan k -számú részrendszere, amelyeknek a gradiensei az $(\mathbf{x}_0, \mathbf{0})$ pontban függetlenek és a többi feltétel pedig az ezen feltételek által meghatározott elemi felületre nézve redundáns az $(\mathbf{x}_0, \mathbf{0})$ pontban.

Bizonyítás. I. Tegyük fel, hogy az $M=T$ feltétel teljesül. Mivel az előbbi definícióból következik, hogy egy redundáns feltétel gradiense benne van a lineárisan független gradiensek által felfeszített altérben, így az M halmaz $(n+k_1-k)$ -dimenziós, ezért a T halmaznak is $(n+k_1-k)$ -dimenziósnak kell lenni. Ebből következik, hogy az elemi felület előállításában szereplő független változók (a felület paramétereinek) száma is $(n+k_1-k)$, tehát azon változók száma, amelyeket a többiek segítségével kifejezhetünk k . Azonban a feltevés szerint van k darab független gradiensű feltétel, ezért az implicit függvény tételt felhasználva k változó kifejezhető a többi segítségével. Mivel az elemi felület egyértelmű, ezért a többi feltételnek redundánsnak kell lennie az elemi felületre nézve.

II. Tegyük fel, hogy a feltételeknek van olyan k számú részrendszere, amelyeknek a gradiensei az $(\mathbf{x}_0, \mathbf{0})$ pontban függetlenek és a többi feltétel pedig ezekre a feltételekre nézve redundáns az $(\mathbf{x}_0, \mathbf{0})$ pontban. Ebből következik, hogy az elemi felületet már a k független gradiensű feltétel meghatározza, ezért az elemi felület paramétereinek száma $(n+k_1-k)$. Mivel az $M \supset T$ reláció teljesül, ez csak úgy lehet, ha az M is $(n+k_1-k)$ -dimenziós altér, ezért $M=T$.

A most következő részben az egyenlőtlenség feltételekkel megadott feladatoknál a szigorú lokális optimalitást biztosító elegendő feltételeket vezetjük le az 1.3. tétel segítségével. Mivel a (2.8) képletből látható, hogy a $\nabla_{(\mathbf{x}, \mathbf{z})}^2 L(\mathbf{x}_0, \mathbf{0}, \boldsymbol{\mu})$ mátrix nem pozitív definit az M halmazon, ha valamelyik i indexre $\mu_i=0$, ezért az 1.3. tétel közvetlenül nem alkalmazható, hanem csak az alábbi állítás bizonyítható.

2.3. TÉTEL. Ha a (2.1) feladatot tekintve léteznek $\mu_i \geq 0$, $i=1, \dots, m$ értékek úgy, hogy

$$(2.10) \quad \nabla f(\mathbf{x}_0) - \sum_{i=1}^m \mu_i \nabla g_i(\mathbf{x}_0) = 0,$$

$$(2.11) \quad \mu_i g_i(\mathbf{x}_0) = 0, \quad i = 1, \dots, m \quad \text{és}$$

$$(2.12) \quad \mathbf{y}^T \nabla^2 f(\mathbf{x}_0) \mathbf{y} - \sum_{i=1}^m \mu_i \mathbf{y}^T \nabla^2 g_i(\mathbf{x}_0) \mathbf{y} > 0, \quad \text{ha}$$

$$\nabla g_i(\mathbf{x}_0) \mathbf{y} = 0, \quad i \in \{i | i \in I(\mathbf{x}_0), \mu_i > 0\},$$

akkor \mathbf{x}_0 a (2.1) szigorú lokális minimuma.

Bizonyítás. Tekintsük a (2.1), illetve a (2.2) feladat helyett az alábbi problémát.

$$(2.13) \quad \min f(\mathbf{x}); \quad h_i(\mathbf{x}, \mathbf{z}) = g_i(\mathbf{x}) - \frac{1}{2} z_i^2 = 0, \quad i \in \{i | i \in I(\mathbf{x}_0), \mu_i > 0\}.$$

Ennek a feladatnak az értelmezési tartománya bővebb mint az eredetie, így ha $(\mathbf{x}_0, \mathbf{0})$ a (2.13) lokális optimuma és \mathbf{x}_0 megengedett, akkor \mathbf{x}_0 lokális optimuma lesz az eredetinek is.

Alkalmazva most erre a problémára az 1.3. tételt, megkapjuk a bizonyítandó állítást.

Megjegyezzük, hogy (2.10), (2.11), (2.12) az optimalitás jól ismert elegendő feltételei [1], [7].

IRODALOM

- [1] AVRIEL, M., *Nonlinear Programming, Analysis and Methods* (Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1976).
- [2] BAZARAA, M. S. and SHETTY, C. M., *Foundations of Optimization* (Springer-Verlag, Berlin, Heidelberg, New York, 1976).
- [3] FIACCO, A. V. and McCORMICK, G. P., *Nonlinear Programming, Sequential Unconstrained Minimization Techniques* (Wiley and Sons, New York, 1968).
- [4] KARUSH, W., Minima of functions of several variables with inequalities as side conditions, Dissertation for the degree of master of science, Chicago, Illinois, 1939.
- [5] KUHN, H. W. and TUCKER, A. W., Nonlinear programming, in Jerzy Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, U. of Calif. Press, 1950), 481—492.
- [6] KUHN, H. W., "Nonlinear programming: a historical view", *Nonlinear Programming*, SIAM—AMS Proceeding Vol. IX. (American Mathematical Society, Providence, Rhode Island, (1976).
- [7] LUENBERGER, D. G., *Introduction to Linear and Nonlinear Programming* (Addison-Wesley Publishing Company Inc. 1973).
- [8] MANGASARIAN, O. L., *Nonlinear Programming* (Mc Graw-Hill Book Company, 1969).
- [9] MARTOS, B., *Nonlinear Programming Theory and Methods* (Akadémiai Kiadó, Budapest, 1975).
- [10] PRÉKOP, A., „Az optimalizáláselmélet kialakulásának történetéről”, *Alkalmazott Matematikai Lapok* 4 (1978) 165—191.
- [11] RAPCSÁK, T., „A nemlineáris programozási feladat optimalitási feltételeinek differenciálgeometriai vizsgálata”, *Alkalmazott Matematikai Lapok* (közlésre elfogadva).
- [12] ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, Inc. Englewood Cliffs, 1969).

(Beérkezett: 1983. március 7.)

RAPCSÁK TAMÁS

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1250 BUDAPEST, ÜRI U. 49.

REMARKS ON MINIMALIZATION PROBLEMS WITH INEQUALITY CONSTRAINTS

T. RAPCSÁK

In this paper it is shown that the first and second-order, necessary and sufficient optimality conditions of the nonlinear programming problem of the inequality type are the generalization of the same conditions for single variable functions.

EGY BIKRITÉRIUM FELADAT MEGOLDÁSÁRÓL

DORMÁNY MIHÁLY

Miskolc

A dolgozat egy speciális többcélú programozási feladattal foglalkozik, amikor is egy korlátos konvex poliéder felett két célfüggvényt kell maximálnunk. Az egyik célfüggvény lineáris (pl. nyereségvolumen), a másik célfüggvény hiperbolikus (pl. a ráfordításegységre jutó fajlagos nyereség). Mivel a két célfüggvény súlyozott összege nem rendelkezik ún. kvázimonotonitási tulajdonsággal, a *simplex módszer* közvetlenül nem alkalmazható, de mégis felhasználást nyer, amikor a célfüggvény nevezőjének paraméterezésével feladatunkat egy duál-paraméteres LP feladatra vezetjük vissza. Alkalmasság normálással a *Pareto-optimalis programok* stabilitását is vizsgálhatjuk.

1. Bevezetés

A többcélú feladatok a matematikai programozásnak olyan problémái, melyek mind elméletileg, mind gyakorlatilag a legérdekesebbek közé tartoznak. Általában nincs ugyanis olyan program, amely minden figyelembe vett cél szempontjából optimalis lenne — a döntés ilyenkor csak valamilyen kompromisszum alapján történhet. A kiválasztott kompromisszumos program — az alkalmazott döntési stratégiától függően — a *Pareto-optimalis* (*efficiens*, *hatékony* néven is ismert) programok közül kerülhet ki. (*Pareto-optimalisnak* az olyan programokat nevezik, melyek a megvalósítható megoldások tartományán nem változtathatók meg anélkül, hogy ezáltal legalább egy figyelembe vett célt le ne rontanánk.)

A többcélú lineáris programozás (másképpen a „lineáris vektor-maximum probléma”) általánosan ismert, jól kimunkált terület (pl. [4], [8]). Vizsgálatok folynak a nemfolytonos lineáris (pl. [2], [9]) és a folytonos nemlineáris modellekkel kapcsolatban is — ez utóbbi esetben inkább csak elvi síkon, pl. a hatékonyság fogalmának szigorítása (*proper efficiency*) irányában ([1], [5]).

2. A feladat megfogalmazása

A dolgozat azt az esetet vizsgálja, amikor az

$$L = \{x | x \geq 0, Ax \leq b\} \subset R^n$$

nemüres korlátos konvex poliéder felett az

$$f_1(x) = c^T x \quad \text{és}$$

$$f_2(x) = \frac{c^T x}{d^T x + \delta}, \quad (d^T x + \delta > 0, \text{ ha } x \in L)$$

célfüggvényeket kell maximálnunk.

A feladat annak a termelésprogramozónak a dilemmáját tükrözi, aki nem tudja eldönteni, hogy a nyereségvolument vagy a ráfordításegységre jutó fajlagos nyereséget maximálja-e. (Ebben az esetben c a fajlagos hozam, d a proporcionális költség vektora, δ a fix költség és L a megvalósítható programok halmaza.)

Ismeretes (pl. [7]), hogy a *Pareto-optimalis programok* meghatározásának problémája visszavezethető egy primál-parametrikus programozási feladat megoldására. Eszerint, ha az L halmaz felett az $f_1(x)$, $f_2(x)$, ..., $f_r(x)$ függvényeket kell egyidejűleg maximálnunk, akkor ezt a problémát a

$$(2.1) \quad \max_{x \in L} \left\{ \sum_{j=1}^r \alpha_j f_j(x) \mid \forall \alpha_j > 0, \sum_{j=1}^r \alpha_j = 1 \right\}$$

feladattal hozhatjuk kapcsolatba. Ugyanis az L halmaznak minden olyan pontja, amely bármely, a feltételeknek eleget tevő α súlyrendszer esetén a (2.1) feladat megoldása, *Pareto-optimalis program*.

Bikritérium feladatunk így — az

$$f(x, \alpha) = \alpha c^T x + (1 - \alpha) \frac{c^T x}{d^T x + \delta} = c^T x \left(\alpha + \frac{1 - \alpha}{d^T x + \delta} \right)$$

jelöléssel — az

$$(2.2) \quad f_{\max}(\alpha) = \max_{x \in L} f(x, \alpha), \quad 0 < \alpha < 1$$

primál-parametrikus feladathoz vezet. α értékét 0-tól 1-ig végigfuttatva *Pareto optimalis pontokat* kapunk.

Mivel esetünkben $f_1(x)$ és $f_2(x)$ mind jelentésében, mind nagyságrendben eltér egymástól, a súlyozott összegben elvész a közgazdasági tartalom. Szokásos eljárás, hogy ilyenkor a két célfüggvényt az $f_{1\max} = \max_{x \in L} f_1(x)$, ill. $f_{2\max} = \max_{x \in L} f_2(x)$ értékkel elosztva „normáljuk”. Különösen indokolt ez a lépés abban az esetben, amikor mind $f_1(x)$, mind pedig $f_2(x)$ az L halmaz felett a zérust veszi fel minimális értéként. Ilyenkor ugyanis bármely $x \in L$ programra az $\frac{f_i(x)}{f_{i\max}}$, $i=1, 2$ érték az illető célfüggvény

„megvalósítási fokaként” értelmezhető. Ebben a felfogásban a többcélú programozási feladat nem más, mint az „ α súlyrendszerrel vett átlagos megvalósítási fok” maximalizálása. Ez az eljárás nagy segítségünkre van akkor is, amikor a gyakorlati megvalósítás céljából a *Pareto-optimalis programok* közül egy „kompromisszumos” megoldást kell kiválasztanunk. Ilyen elv lehet pl. annak a programnak a kiválasztása, amely a legnagyobb α intervallum felett optimalis.

A dolgozat a (2.2) feladat megoldására parametrikus eljárást (nevezetesen: duál-parametrikus LP feladatra való visszavezetést) mutat be. Ugyanis, bár mind $f_1(x)$, mind pedig $f_2(x)$ kvázimonoton függvény, könnyen belátható, hogy $f(x, \alpha)$ még rögzített α esetén sem az, így a simplex módszer közvetlen alkalmazhatósága nem biztosított. (A kvázimonotonitás definícióját és az ide vonatkozó tételt lásd pl. [6], 55. és 94. o.)

3. A feladat megoldása

A (2.2) feladat megoldásához szükséges átalakítások az alábbiak (a (3.1)–(3.4) lépésben használt jelölések és összefüggések magyarázatát a (3.4) lépés után fogjuk megadni):

$$(3.1) \quad f_{\max}(\alpha) = \max_{\mathbf{x} \in L} f(\mathbf{x}, \alpha) = \max_{\mathbf{x} \in L} \left\{ \mathbf{c}^T \mathbf{x} \left(\alpha + \frac{1-\alpha}{\mathbf{d}^T \mathbf{x} + \delta} \right) \right\}$$

$$= \max_{t \in T} \max_{\mathbf{x} \in L'(t)} \left\{ \mathbf{c}^T \mathbf{x} \left(\alpha + \frac{1-\alpha}{t} \right) \right\}$$

$$(3.2) \quad = \max_{t \in T} \left\{ \left(\alpha + \frac{1-\alpha}{t} \right) \max_{\mathbf{x} \in L'(t)} \mathbf{c}^T \mathbf{x} \right\}$$

$$(3.3) \quad = \max_{t \in T} \left\{ \left(\alpha + \frac{1-\alpha}{t} \right) \mathcal{F}(t) \right\}$$

$$(3.4) \quad = \max_{1 \leq k \leq p} \max_{t \in T_k} \left\{ \left(\alpha + \frac{1-\alpha}{t} \right) (g_k t + h_k) \right\}.$$

ad (3.1): A $\mathbf{d}^T \mathbf{x} + \delta = t$ transzformációval vezessük be a t paramétert. (Mivel L felett $\mathbf{d}^T \mathbf{x} + \delta > 0$ kell legyen, ezért $t > 0$). Legyen

$$L'(t) = L \cap \{\mathbf{x} | \mathbf{d}^T \mathbf{x} + \delta = t\};$$

és

$$T = \{t | t = \mathbf{d}^T \mathbf{x} + \delta, \mathbf{x} \in L\},$$

a paraméter értelmezési tartománya. (Egyparaméteres esetről lévén szó, T nyilvánvalóan egy szakasz.)

ad (3.2): A gömbölyű zárójelben levő rész nem függ az \mathbf{x} változótól, minthogy t rögzített.

ad (3.3): Legyen

$$(3.5) \quad \mathcal{F}(t) = \max_{\mathbf{x} \in L'(t)} \mathbf{c}^T \mathbf{x} = \max \{ \mathbf{c}^T \mathbf{x} | \mathbf{x} \geq 0, \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{d}^T \mathbf{x} + \delta = t \}$$

a duál-parametrikus LP feladat megoldása.

ad (3.4): Mint ismeretes (pl. [3]), $\mathcal{F}(t)$ szakaszonként (karakterisztikus tartományonként) lineáris folytonos konkáv függvény, ahol ezek a karakterisztikus tartományok az egész T tartományt lefedik; t értékeinek vizsgálatánál azonban elegendő csupán egy később meghatározandó $[t_{\min}, t_{\max}]$ intervallum vizsgálata. ($t_{\min} = t_{\max}$ esetben — mint látni fogjuk — további vizsgálatra nincs szükség.) t_{\min} , ill. t_{\max} meghatározásánál abból fogunk kiindulni, hogy $t < t_{\min}$, ill. $t > t_{\max}$ esetben egyik függ-

vény értéke sem nő (és legalább az egyiké csökken). Legyen

$$\mathcal{F}(t) \supset F(t) = \begin{cases} F_1(t) = g_1 t + h_1, & \text{ha } t \in T_1 = [t_1, t_2], \\ \vdots \\ F_k(t) = g_k t + h_k, & \text{ha } t \in T_k = [t_k, t_{k+1}], \\ \vdots \\ F_p(t) = g_p t + h_p, & \text{ha } t \in T_p = [t_p, t_{p+1}], \end{cases}$$

ahol $t_1 = t_{\min}$, $t_{p+1} = t_{\max}$.

Legyen \mathbf{x}_1^0 , ill. \mathbf{x}_2^0 az $f_1(\mathbf{x})$, ill. $f_2(\mathbf{x})$ függvénynek olyan maximumhelye, amely egyúttal *Pareto-optimális program* is. Ha mindkét függvény egyetlen maximumhellyel rendelkezik az L halmazon, akkor nincs probléma; alternatív optimumok esetén azonban némi körütekintésre van szükség. Ugyanis az alternatív optimumhelyek közül csak azok a programok lehetnek *Pareto-optimálisak*, melyek a *másik célfüggvényre nézve* dominálják a többi optimumhelyet. Vagyis, ha \mathbf{x}_1^i , $i \in I_0 = \{1, 2, \dots, i_0\}$, ill. \mathbf{x}_2^j , $j \in J_0 = \{1, 2, \dots, j_0\}$ jelöli az L halmaznak azon csúcseit, melyek az első, ill. második feladatra nézve optimálisak, akkor — a := jelet az optimumhely jelölésére használva —

$$\mathbf{x}_2^0 := \max_{j \in J_0} \mathbf{c}^T \mathbf{x}_2^j, \quad \text{és} \quad \mathbf{x}_1^0 := \max_{i \in I_0} f_2(\mathbf{x}_1^i).$$

Mivel az $f_2(\mathbf{x})$ törtfüggvény nevezője mindig pozitív, ezért az utóbbi összefüggéssel egyenértékűen

$$\mathbf{x}_1^0 := \begin{cases} \min_{i \in I_0} \mathbf{d}^T \mathbf{x}_1^i, & \text{ha } f_{1\max} > 0, \\ \max_{i \in I_0} \mathbf{d}^T \mathbf{x}_1^i, & \text{ha } f_{1\max} < 0. \end{cases}$$

\mathbf{x}_1^0 és \mathbf{x}_2^0 nem feltétlenül egyértelműen meghatározott, akárcsak az $f_{1\max} = f_{2\max} = 0$ esetben, amikor is az alternatív optimális csúcok halmaza mindkét feladatra nézve azonosan $\{\mathbf{x}_1^i\} \equiv \{\mathbf{x}_2^j\}$; \mathbf{x}_1^0 , ill. \mathbf{x}_2^0 pedig ezen halmaz bármely eleme.

\mathbf{x}_1^0 és \mathbf{x}_2^0 fenti értelmezéséből következik, hogy

$$\mathbf{c}^T \mathbf{x}_2^0 \leq \mathbf{c}^T \mathbf{x}_1^0$$

és

$$\frac{\mathbf{c}^T \mathbf{x}_1^0}{\mathbf{d}^T \mathbf{x}_1^0 + \delta} \leq \frac{\mathbf{c}^T \mathbf{x}_2^0}{\mathbf{d}^T \mathbf{x}_2^0 + \delta}.$$

E két egyenlőtlenség összekapcsolásából

$$(3.6) \quad \frac{\mathbf{c}^T \mathbf{x}_2^0}{\mathbf{d}^T \mathbf{x}_1^0 + \delta} \leq \frac{\mathbf{c}^T \mathbf{x}_1^0}{\mathbf{d}^T \mathbf{x}_1^0 + \delta} \leq \frac{\mathbf{c}^T \mathbf{x}_2^0}{\mathbf{d}^T \mathbf{x}_2^0 + \delta}$$

következik, innen pedig — mivel valamennyi nevező pozitív — az alábbi 3 lehetséges eset:

$$(I) \quad \mathbf{c}^T \mathbf{x}_2^0 = \mathbf{c}^T \mathbf{x}_1^0,$$

$$(II) \quad 0 < \mathbf{c}^T \mathbf{x}_2^0 < \mathbf{c}^T \mathbf{x}_1^0,$$

$$(III) \quad \mathbf{c}^T \mathbf{x}_2^0 < \mathbf{c}^T \mathbf{x}_1^0 < 0.$$

ad (I): Ha \mathbf{x}_1^0 , ill \mathbf{x}_2^0 egyértelműen meghatározott, akkor $\mathbf{x}_1^0 = \mathbf{x}_2^0$ az egyetlen *Pareto-optimum*; ha nem egyértelműen meghatározott, akkor \mathbf{x}_1^0 és \mathbf{x}_2^0 definíciójából következően ezek a pontok a két feladat közös alternatív optimumai, így valamennyien (és a konvex kombinációk is) *Pareto-optimálisak*.

Ebben az esetben tulajdonképpen nem is beszélhetünk bikritérium problémáról; a feladat megoldása csupán az esetleges alternatív optimumok vizsgálatára korlátozódik.

ad (II): Ez az eset a gyakorlat számára a legérdekesebb, és — mint látni fogjuk — könnyebben kezelhető, mint (III). (3.6)-ból következően $\mathbf{d}^T \mathbf{x}_2^0 + \delta < \mathbf{d}^T \mathbf{x}_1^0 + \delta$, vagyis

$$t_{\min} = \mathbf{d}^T \mathbf{x}_2^0 + \delta, \quad t_{\max} = \mathbf{d}^T \mathbf{x}_1^0 + \delta.$$

ad (III): (3.6)-ból következően most $\mathbf{d}^T \mathbf{x}_1^0 + \delta < \mathbf{d}^T \mathbf{x}_2^0 + \delta$, vagyis $t_{\min} = \mathbf{d}^T \mathbf{x}_1^0 + \delta$, $t_{\max} = \mathbf{d}^T \mathbf{x}_2^0 + \delta$.

A (II) és (III) eset megkülönböztetése g_k értékeinek a (3.4) összefüggésben felvett előjele szempontjából is érdekes. A (3.5) feladatból nyilvánvaló ugyanis, hogy $f_1(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$ csökkenése $\mathcal{F}(t)$ csökkenését is maga után vonja, vagyis $\mathcal{F}(t)$ — és így $F(t)$ is — a vizsgálandó $[t_{\min}, t_{\max}]$ intervallum egyik végpontjában — aholis $t = \mathbf{d}^T \mathbf{x}_1^0 + \delta$ — veszi fel maximális értékét. Ebből — és $F(t)$ konkávitásából — viszont az következik, hogy a (3.4) összefüggésben a (II) esetben valamennyi $g_k \geq 0$, a (III) esetben pedig valamennyi $g_k \leq 0$. A g_k együtthatók előjelét azért kell figyelnünk, mert rögzített $0 < \alpha < 1$ esetén a

$$(3.7) \quad H_k(t) = \left(\alpha + \frac{1-\alpha}{t} \right) (g_k t + h_k)$$

egyváltozós függvény (hiperbola) szabad lokális maximuma

$$(A) \quad g_k > 0 \text{ (és } h_k \geq 0) \text{ esetén a } \tau_k(\alpha) = -\sqrt{\frac{1-\alpha}{\alpha} \frac{h_k}{g_k}} \leq 0 \text{ helyen,}$$

$$(B) \quad g_k < 0 \text{ (és } h_k \leq 0) \text{ esetén a } \tau_k(\alpha) = \sqrt{\frac{1-\alpha}{\alpha} \frac{h_k}{g_k}} \geq 0$$

helyen van. (Egyéb esetekben — vagyis, ha $g_k = 0$, vagy g_k és h_k előjele különbözik — nincs szabad lokális maximum.)

Mivel megállapodásunk értelmében az L halmaz felett $t > 0$, ezért minden olyan esetben, amikor nincs szabad lokális maximum, vagy az (A) esetnek megfelelően nempozitív τ_k helyen van, (3.4)-ben a t szerinti maximálást elegendő csupán T_k végpontjaira elvégezni.

(B) esetben (ezen belül is, ha $h_k < 0$), a $T_k = [t_k, t_{k+1}]$ intervallumban lokális maximum van a $\tau_k(\alpha)$ helyen, ha

$$t_k \leq \sqrt{\frac{1-\alpha}{\alpha} \frac{h_k}{g_k}} \leq t_{k+1},$$

vagyis, ha

$$\alpha'_k \leq \alpha \leq \alpha''_k,$$

ahol

$$\alpha'_k = \frac{1}{1 + \frac{g_k}{h_k} t_{k+1}^2}, \quad \alpha''_k = \frac{1}{1 + \frac{g_k}{h_k} t_k^2}.$$

Mivel a T_1, \dots, T_p szakaszok szomszédoként közös határponttal rendelkeznek, és $F(t)$ folytonos függvény, ezért a (B) eset kivételével elegendő pl. minden szakasz alsó határpontját — hozzávéve a p -edik szakasz felső határpontját, aholis $t_{p+1} = t_{\max}$ — vizsgálni. (Így pl. a T_k tartományhoz a

$$(3.8) \quad H_k(\alpha) = \left(\alpha + \frac{1-\alpha}{t_k} \right) (g_k t_k + h_k)$$

egyenest rendelkezhetjük.)

A helyzet a (B) esetben annyival bonyolultabb, hogy a T_k tartományhoz hozzárendelhető $H_k(\alpha)$ függvény a $0 < \alpha \leq \alpha'_k$ és $\alpha''_k \leq \alpha < 1$ szakaszon megegyezik a (3.8) egyenessel, de az $\alpha'_k \leq \alpha \leq \alpha''_k$ intervallumon — a (3.7) függvényt a $t = \tau_k(\alpha)$ helyen véve — a

$$H_k(\alpha) = \sqrt{4\alpha(1-\alpha)g_k h_k} + \alpha(h_k - g_k) + g_k$$

függvénynek (ellipszis) megfelelően „kidudorodik”. Vagyis összefoglalóan legyen

$$H_k(\alpha) = \begin{cases} \sqrt{4\alpha(1-\alpha)g_k h_k} + \alpha(h_k - g_k) + g_k, & \text{ha } g_k, h_k < 0 \text{ és } \alpha'_k \leq \alpha \leq \alpha''_k, \\ \left(\alpha + \frac{1-\alpha}{t_k} \right) (g_k t_k + h_k) & \text{egyébként,} \end{cases}$$

$$k = 1, 2, \dots, p \text{ és } \dots$$

$$H_{p+1}(\alpha) = \left(\alpha + \frac{1-\alpha}{t_{\max}} \right) (g_p t_{\max} + h_p).$$

Az $f_{\max}(\alpha)$ ezen $H_k(\alpha)$, $k=1, 2, \dots, p, p+1$ görbesereg (a gyakorlati problémák-nál általában egyenessereg) felső burkolója, vagyis

$$f_{\max}(\alpha) = \max_{1 \leq k \leq p+1} H_k(\alpha), \quad 0 < \alpha < 1.$$

Jelölje $\mathbf{x}^k(t)$ a T_k tartomány felett optimális bázismegoldást, vagyis a (3.5) feladat \mathbf{x} szerinti optimumhelyét, midőn $t_k \leq t \leq t_{k+1}$. Jelölje továbbá $\alpha_0=0, \alpha_1, \dots, \alpha_{s-1}, \alpha_s, \dots, \alpha_q=1$ az $f_{\max}(\alpha)$ felső burkoló töréspontjait abban az értelemben, hogy a burkoló α_{s-1} értéknél lép valamely $H_r(\alpha)$ görbére, és α_s értéknél lép le róla, vagyis

$$f_{\max}(\alpha) = H_r(\alpha), \quad \text{ha } \alpha \in (\alpha_{s-1}, \alpha_s); \quad 1 \leq r \leq p+1.$$

Ekkor a *Pareto-optimális program(ok)*

— a (B) esetben

az $(\alpha_{s-1}, \alpha_s) \cap (\alpha'_r, \alpha''_r)$ szakaszon $\mathbf{x}^r(\tau_r(\alpha))$,

az $(\alpha_{s-1}, \alpha_s) \setminus (\alpha'_r, \alpha''_r)$ szakasz(ok)on $\mathbf{x}^r(t_r)$,

— egyébként az (α_{s-1}, α_s) szakaszon végig $\mathbf{x}^r(t_r)$.

Az α_s , $1 \leq s \leq q-1$ töréspontban — a bal oldali $H_s(\alpha)$, és a jobb oldali, mondjuk $H_t(\alpha)$ görbének megfelelően — várhatóan két optimális bázismegoldás is adódik. (Ugyanis itt az $f(x, \alpha_s)$ függvénynek alternatív optimumai lehetnek.)

Végezetül érdemes megemlíteni, hogy az $x^k(t_k)$ bázismegoldások az L halmaz csúcsai. Ugyanis — ha az A mátrix sorainak száma m — a (3.5) feladatban, a $d^T x + \delta = t$ újabb feltétel miatt a bázisváltozók száma $m+1$. A T_k tartomány t_k határpontjában viszont az $x^k(t_k)$ bázismegoldás *degenerált*, vagyis a nemzéró komponensek száma legfeljebb m — ami jelzi, hogy a pont extrémális.

Az algoritmus főbb lépései a következők:

(1) Legyen $L = \{x | x \geq 0, Ax \leq b\}$, és legyen

$$f_{1\max} = \max_{x \in L} c^T x,$$

$$f_{2\max} = \max_{x \in L} \frac{c^T x}{d^T x + \delta}.$$

Legyenek az első, ill. második feladat esetleges alternatív optimális bázismegoldásai x_1^i , $i \in I_0$, ill. x_2^j , $j \in J_0$.

Ha $f_{1\max} > 0$, legyen $x_1^0 := \min_{i \in I_0} d^T x_1^i$ és $x_2^0 := \max_{j \in J_0} c^T x_2^j$,

ha $f_{1\max} < 0$, legyen $x_1^0 := \max_{i \in I_0} d^T x_1^i$ és $x_2^0 := \max_{j \in J_0} c^T x_2^j$.

(2) Ha $f_{1\max} = 0$ vagy $c^T x_1^0 = c^T x_2^0$, ugorjunk (9)-re.

Egyébként

(3) $f_{1\max} > 0$ esetén legyen $t_{\min} = d^T x_2^0 + \delta$, $t_{\max} = d^T x_1^0 + \delta$,

$f_{1\max} < 0$ esetén legyen $t_{\min} = d^T x_1^0 + \delta$, $t_{\max} = d^T x_2^0 + \delta$.

(4) Oldjuk meg a

$$\max \{c^T x | x \geq 0, Ax \leq b, d^T x + \delta = t, t_{\min} \leq t \leq t_{\max}\}$$

duál-parametrikus LP feladatot. A megoldás legyen

$$F(t) = g_k t + h_k, \quad \text{ha } t \in [t_k, t_{k+1}], \quad k = 1, 2, \dots, p.$$

(5) $k = 1$.

(6) Ha $g_k < 0$ és $h_k < 0$, legyen

$$\tau_k(\alpha) = \sqrt{\frac{1-\alpha}{\alpha} \frac{h_k}{g_k}},$$

$$\alpha'_k = \frac{h_k}{h_k + g_k t_{k+1}^2}, \quad \alpha''_k = \frac{h_k}{h_k + g_k t_k^2};$$

egyébként legyen $\alpha'_k = \alpha''_k = -1$.

$$(7) \quad H_k(\alpha) = \begin{cases} \sqrt{4\alpha(1-\alpha)g_k h_k} + \alpha(h_k - g_k) + g_k, & \alpha \in (\alpha'_k, \alpha''_k), \\ \left(\alpha + \frac{1-\alpha}{t_k}\right)(g_k t_k + h_k) & \text{egyébként.} \end{cases}$$

(8) Ha $k < p$, akkor $k \leftarrow k+1$ és térjünk vissza (6)-ra. Ha $k = p$, akkor legyen

$$H_{p+1}(\alpha) = \left(\alpha + \frac{1-\alpha}{t_{\max}}\right)(g_p t_{\max} + h_p).$$

Ugorjunk (10)-re.

(9) $\mathbf{c}^T \mathbf{x}_1^0 = \mathbf{c}^T \mathbf{x}_2^0$ esetén a *Pareto-optimális program(ok)* a mindkét feladatra nézve optimális csúcspont(ok) konvex kombinációi).

Ugorjunk (11)-re.

(10) A felső burkoló:

$$f_{\max}(\alpha) = \max_{1 \leq k \leq p+1} H_k(\alpha).$$

Jelölje $\mathbf{x}^k(t)$ a T_k tartomány felett optimális bázismegoldást, és legyenek $\alpha_0 = 0$, $\alpha_1, \dots, \alpha_{s-1}$, $\alpha_s, \dots, \alpha_q = 1$ a felső burkoló töréspontjai úgy, hogy $f_{\max}(\alpha) = H_r(\alpha)$, $\alpha \in (\alpha_{s-1}, \alpha_s)$. A *Pareto-optimális programok*:

$$\mathbf{x}^r(\tau_r(\alpha)), \quad \alpha \in (\alpha_{s-1}, \alpha_s) \cap (\alpha'_r, \alpha''_r),$$

$$\mathbf{x}^r(t_r), \quad \alpha \in (\alpha_{s-1}, \alpha_s) \setminus (\alpha'_r, \alpha''_r).$$

Az α_s , $1 \leq s \leq q-1$ töréspontokban alternatív optimumokkal kell számolnunk.

(11) VÉGE.

4. Mintapélda

Legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 & 3 \\ 3 & 1 & 2 & 1 \\ 2 & 3 & 3 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 40 \\ 40 \\ 46 \end{bmatrix},$$

$$\mathbf{c} = [4, 5, 8, 5]^T, \quad \mathbf{d} = [4, 5, 5, 1]^T, \quad \delta = 100.$$

Az induló szimplex táblát olyan alakban vesszük fel, amely mind $f_1(\mathbf{x})$, mind $f_2(\mathbf{x})$ maximálására alkalmas. A hiperbolikus programozási feladatot MARTOS módszerével (lásd pl. [6], 176. o.) oldjuk meg. Induláskor $\mathbf{t}(0) = \mathbf{d}^T \mathbf{0} + \delta = 100$.

1. TÁBLÁZAT

	x_1	x_2	x_3	x_4	
v_1	1	2	4	3	40
v_2	3	1	2	1	40
v_3	2	3	3	2	46
$-f_1$	4	5	8	5	0
$-t$	4	5	5	1	-100

$f_2(\mathbf{x})$ maximumát tartalmazó transzformált táblázat (egyetlen optimális megoldás létezik):

2. TÁBLÁZAT

	v_2	x_2	x_3	v_1	
x_4	$-\frac{1}{8}$	$\frac{5}{8}$	$\frac{5}{4}$	$\frac{3}{8}$	10
x_1	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$-\frac{1}{8}$	10
v_3	$-\frac{1}{2}$	$\frac{3}{2}$	0	$-\frac{1}{2}$	6
$-f_1$	$-\frac{7}{8}$	$\frac{11}{8}$	$\frac{3}{4}$	$-\frac{11}{8}$	-90
$-t$	$-\frac{11}{8}$	$\frac{31}{8}$	$\frac{11}{4}$	$\frac{1}{8}$	-150

$f_1(\mathbf{x})$ maximumát tartalmazó transzformált táblázat:

3. TÁBLÁZAT

	v_2	v_3	v_1	x_4	
x_3	$\frac{1}{15}$	$-\frac{1}{3}$	$\frac{7}{15}$	$\frac{4}{5}$	6
x_1	$\frac{2}{5}$	0	$-\frac{1}{5}$	$-\frac{1}{5}$	8
x_2	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	0	4
$-f_1$	$-\frac{7}{15}$	$-\frac{2}{3}$	$-\frac{19}{15}$	$-\frac{3}{5}$	-100
$-t$	$-\frac{4}{15}$	$-\frac{5}{3}$	$\frac{2}{15}$	$-\frac{11}{5}$	-182

Mivel $f_1(\mathbf{x}_2^0) = 90 < 100 = f_{1\max}$, ezért a duál-parametrikus LP feladat megoldása következik. A szimplex táblák utolsó sora tartalmazza a $\mathbf{d}^T \mathbf{x} + \delta = t$ egyenlet transzformált alakját, mely egyenletet — célszerűen a 2. táblázatot véve — a feltételek közé sorolunk (így ugyanis rögtön a $t_1 = t_{\min}$ értékről indulhatunk):

4. TÁBLÁZAT

	v_2	x_2	x_3	v_1	
x_4	$-\frac{1}{8}$	$\frac{5}{8}$	$\frac{5}{4}$	$\frac{3}{8}$	10
x_1	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$-\frac{1}{8}$	10
v_3	$-\frac{1}{2}$	$\frac{3}{2}$	0	$-\frac{1}{2}$	6
*	$-\frac{11}{8}$	$\frac{31}{8}$	$\frac{11}{4}$	$\frac{1}{8}$	$-150 + t$
$-F(t)$	$-\frac{7}{8}$	$\frac{11}{8}$	$\frac{3}{4}$	$-\frac{11}{8}$	-90

A bekeretezett helyen választva generáló elemet, a duál-parametrikus LP feladat első optimális táblájához jutunk:

5. TÁBLÁZAT

	v_2	x_3	v_1	
x_4	$\frac{3}{31}$	$\frac{25}{31}$	$\frac{11}{31}$	$\frac{1060}{31} - \frac{5}{31}t$
x_1	$\frac{13}{31}$	$\frac{5}{31}$	$\frac{4}{31}$	$\frac{460}{32} - \frac{1}{31}t$
v_2	$\frac{1}{31}$	$\frac{33}{31}$	$\frac{17}{31}$	$\frac{1986}{31} - \frac{12}{31}t$
x_2	$-\frac{11}{31}$	$\frac{22}{31}$	$\frac{1}{31}$	$-\frac{1200}{31} + \frac{8}{31}t$
$-F(t)$	$-\frac{16}{31}$	$-\frac{7}{31}$	$-\frac{44}{31}$	$-\frac{1140}{31} - \frac{11}{31}t$

Újabb transzformációval jutunk a második — és egyben utolsó — optimális táblához:

6. TÁBLÁZAT

	v_2	v_3	v_1	
x_4	$\frac{4}{33}$	$\frac{25}{33}$	$-\frac{2}{33}$	$\frac{910}{11} - \frac{5}{11}t$
x_1	$\frac{14}{33}$	$\frac{5}{33}$	$-\frac{7}{33}$	$\frac{270}{11} - \frac{1}{11}t$
x_3	$-\frac{1}{33}$	$-\frac{31}{33}$	$\frac{17}{33}$	$-\frac{662}{11} + \frac{4}{11}t$
x_2	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	4
$-F(t)$	$-\frac{13}{33}$	$-\frac{7}{33}$	$-\frac{43}{33}$	$-\frac{554}{11} + \frac{3}{11}t$

A 2. táblázatból leolvasható $t_1 = t_{\min} = 150$, a minimális vizsgálandó paraméter-érték, és $f_{2\max} = \frac{f_1}{t} = \frac{90}{150} = 0,6$.

A 3. táblázatból leolvasható $t_{\max} = 182$, a maximális vizsgálandó paraméter-érték, és $f_{1\max} = 100$.

Az 5. táblázat jobb oldalának nemnegativitása határozza meg a

$$T_1 = \{t | 150 \leq t \leq 165,5\}$$

karakterisztikus tartományt, amely felett

$$F(t) = F_1(t) = \frac{1140}{31} + \frac{11}{31}t.$$

A 6. táblázat jobb oldalának nemnegativitása határozza meg a

$$T_2 = \{t | 165,5 \leq t \leq 182\}$$

karakterisztikus tartományt, amely felett

$$F(t) = F_2(t) = \frac{554}{11} + \frac{3}{11}t.$$

Az utolsó lépés — vagyis $f_{\max}(\alpha)$ meghatározása — előtt végezzük el a 2. fejezetben jelzett normálást, így ugyanis a *Pareto-optimalis programok* meghatározásán túlmenően egy kompromisszumos megoldást is tudunk javasolni. Legyen tehát

$$H_k^*(\alpha) = \begin{cases} \left(\frac{\alpha}{f_{1\max}} + \frac{1-\alpha}{f_{2\max}} \frac{1}{t_k} \right) F_k(t_k), & k = 1, 2, \\ \left(\frac{\alpha}{f_{1\max}} + \frac{1-\alpha}{f_{2\max}} \frac{1}{t_{\max}} \right) F_2(t_{\max}), & k = 3, \end{cases}$$

vagyis

$$H_k^*(\alpha) = \begin{cases} \left(\frac{\alpha}{100} + \frac{1-\alpha}{0,6} \frac{1}{150} \right) 90 = 1 - 0,1\alpha, & k = 1, \\ \left(\frac{\alpha}{100} + \frac{1-\alpha}{0,6} \frac{1}{165,5} \right) 95,5 = 0,9617 - 0,0067\alpha, & k = 2, \\ \left(\frac{\alpha}{100} + \frac{1-\alpha}{0,6} \frac{1}{182} \right) 100 = 0,9158 + 0,0842\alpha, & k = 3. \end{cases}$$

A felső burkoló:

$$f_{\max}^*(\alpha) = \max_{1 \leq k \leq 3} H_k^*(\alpha) = \begin{cases} 1 - 0,1\alpha, & 0 < \alpha \leq 0,4103, \\ 0,9617 - 0,0067\alpha, & 0,4103 \leq \alpha \leq 0,5054, \\ 0,9158 + 0,0842\alpha, & 0,5054 \leq \alpha < 1, \end{cases}$$

és a megfelelő *Pareto-optimalis csúcspontok*

az 5. táblázatból $x^1(t_1) = x^1(150) = [10; 0; 0; 10]^T$,

a 6. táblázatból $x^2(t_2) = x^2(165,5) = [9,5; 4; 0; 7,5]^T$ és

$x^2(t_{\max}) = x^2(182) = [8; 4; 6; 0]^T$.

Mivel a felső burkoló legszélesebb karakterisztikus intervalluma $k=3$ esetre adódik, ezért a $[8; 4; 6; 0]$ pontot tekinthetjük a feladat *kompromisszumos* megoldásának.

Végezetül szeretném megköszönni HOFFER JÁNOS-nak, hogy a kézirat hiányosságaira felhívta a figyelmemet, és ezzel egy teljesebb kidolgozásra ösztönzött.

IRODALOM

- [1] BENSON, H. P. and MORIN, T. L., "The vector maximization problem: proper efficiency and stability", *SIAM Journal on Applied Mathematics* 32 (1977) 64—72.
- [2] BITRAN, G. R., "Theory and algorithms for linear multiple objective programs with zero-one variables", *Mathematical Programming* 17 (1979) 362—390.
- [3] DINKELBACH, W., *Sensitivitätsanalysen und parametrische Programmierung* (Springer, 1969).
- [4] EVANS, J. P. and STEUER, R. E., "A revised simplex method for linear multiple objective programs", *Mathematical Programming* 5 (1973) 54—72.
- [5] GEOFFRION, A. M., "Proper efficiency and the theory of vector maximization", *Journal of Mathematical Analysis and Applications* 22 (1968) 618—630.
- [6] MARTOS, B., *Nonlinear programming* (Akadémiai Kiadó, Budapest, 1975).
- [7] WENDELL, R. E. and LEE, D. N., "Efficiency in multiple objective optimization problems", *Mathematical Programming* 12 (1977) 406—414.
- [8] ZELENÉY, M., *Linear Multiobjective Programming* (Springer, 1974).
- [9] ZIONTS, S., "Integer linear programming with multiple objectives", *Annals of Discrete Mathematics* 1 (1977) 551—562.

(Beérkezett: 1982. március 22.)

DORMÁNY MIHÁLY
NME MATEMATIKAI INTÉZET
MISKOLC-EGYETEMVÁROS, 3515

ON SOLVING A BICRITERION PROGRAM

M. DORMÁNY

The paper deals with a special multiple objective programming problem when two objectives are to be maximized over a bounded konvex polyhedron. One objective is linear, the other is a hyperbolic one. The simplex algorithm is applied when the problem is reduced to a dual-parametric linear programming.

A KRIGING MÓDSZER ALKALMAZÁSA NEMNEGATÍV SÚLYOKKAL ÉS KORLÁTOKKAL

MOLNÁR SÁNDOR és SZIDAROVSKY FERENC

Budapest

A *Kriging módszer* alapvető fontosságú az alkalmazott tudományok szempontjából. Például a bányászat és földtudományok területén az ásványvagyon mennyiségi és minőségi paramétereinek a megbecslése történik a *Kriging módszer* alkalmas változatainak a felhasználásával. SCHAAP és ST. GEORGE kimutatták, hogy a módszer alkalmazásakor nyert súlyok nem feltétlenül adódnak nemnegatívaknak, ez pedig a gyakorlati alkalmazást nehezíti. Ebben a dolgozatban az ő módszerüket továbbfejlesztve kimutatjuk, hogy a súlyokra tett nemnegativitási feltételek mellett a *Kriging módszer* egy alkalmas nemlineáris egyenletrendszer megoldására redukálódik, amely megoldására egy véges algoritmust tudunk javasolni.

1. Bevezetés

A geostatistikai módszerek alapjai az ötvenes évek végén és a hatvanas évek elején kerültek kidolgozásra, amikor MATHERON (1957, 1963, 1969) kidolgozta a regionalizált valószínűségi változók elméletét és annak lehetséges bányászati alkalmazásait. Ezeket a módszereket *Kriging*-nek nevezte el D. G. KRIGE (1951) nyomán, aki először vetette fel hasonló módszerek alkalmazásának igényét és javasolt egy egyszerű módszert, amelyet a *Kriging-módszer* első változatának is tekinthetünk. A bányászati alkalmazások részletes leírását találhatjuk meg KAPOLYI (1981), JOURNEL és HUIJBREGTS (1978) és DAVID (1977) könyvében. A hidrológiai alkalmazások nagyrészt DELHOMME (1978), GAMBOLATI és VOLPI (1979) nyomán történnek.

A „regionalizált valószínűségi változó” elnevezés is MATHERONTól ered és ezen olyan valószínűségi változót értünk általában, amely kiterjed térben és esetleg időben is és speciális struktúrális tulajdonságoknak tesz eleget. Matematikailag, egy $Z(x)$ függvénnyel jellemezhető, ahol x egy, két vagy három dimenziós vektor, és $Z(x)$ jelenti a leírandó paraméter értéket az x helyen. Például, ásványvagyon vastagsága, minőségi paraméterek adott helyeken való értéke, ezek blokkokhoz tartozó átlagértéke mind regionalizált valószínűségi változókként kezelhetők.

A módszerek konkrét leírása megtalálható a MOLNÁR, SZIDAROVSKY, YAKOWITZ (1982) dolgozatban, így annak megismétlésétől eltekintünk.

2. A matematikai modell

Jelölje x_1, \dots, x_n a mérési helyeket, Z_1, \dots, Z_n a megfelelő függvényértékeket, $\gamma(h)$ pedig a variogramot. Tegyük fel továbbá, hogy a Z függvény egy V tartományhoz tartozó Z_v átlagértékét a lineáris

$$(2.1) \quad Z^* = \sum_{i=1}^n \lambda_i Z_i$$

kifejezéssel becsüljük. Ekkor — mint ismeretes — a minimális szórású, torzítatlan becslés együtthatóit a

$$(2.2) \quad \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_{v1} \\ \vdots \\ \gamma_{vn} \\ 1 \end{pmatrix}$$

egyenletrendszer megoldásával kaphatjuk meg, ahol

$$(2.3) \quad \gamma_{ij} = \gamma(x_i - x_j), \quad \gamma_{vi} = \frac{1}{|V|} \int_V \gamma(x - x_i) dx,$$

továbbá a becslés varianciája a

$$(2.4) \quad \text{Var}(Z^*) = -\gamma_{vv} + 2 \sum_{i=1}^n \lambda_i \gamma_{vi} - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma_{ij}$$

képlettel adódik, ahol a λ_i együtthatók a (2.2) egyenletrendszer együtthatóiból adódnak, valamint

$$(2.5) \quad \gamma_{vv} = \frac{1}{|V|^2} \iint_{V \times V} \gamma(x - x') dx dx'.$$

A gyakorlati alkalmazások során gyakran adódnak negatív súlyok, azaz néhány λ_i együttható értéke negatív. SCHAAP és ST. GEORGE (1981) rámutattak, hogy számos alkalmazás esetén a λ_i súlyoknak feltétlenül nemnegatívnak kell lenniük. Ennek illusztrálására tekintsük a következő példát.

2.1. PÉLDA. Tegyük fel, hogy $n=6$, valamint a mérési pontok a következők: $x_1=(0,20)$, $x_2=(20,20)$, $x_3=(0,60)$, $x_4=(0,70)$, $x_5=(0,0)$ és $x_6=(0, -40)$. Tegyük fel továbbá, hogy V egyetlen pontból áll: $V=(10,35)$, és a variogram-függvény a következő:

$$\gamma(h) = \begin{cases} C_0 + C \left[\frac{3}{2} \frac{h}{a} - \frac{1}{2} \frac{h^3}{a^3} \right], & \text{ha } 0 \leq h \leq a \\ C, & \text{ha } h > a, \end{cases}$$

ahol $C_0=0$, $C=1$, $a=80$. Ekkor a (2.2) egyenletrendszer megoldásából a következő súlyok adódnak:

$\lambda_1=0,382$; $\lambda_2=0,403$; $\lambda_3=0,326$; $\lambda_4=-0,013$; $\lambda_5=-0,085$ és $\lambda_6=-0,013$, vagyis λ_4 , λ_5 és λ_6 negatív értékű. Ha a mért függvényértékek a következők lennének:

$$Z_1 = Z_2 = Z_3 = 0 \quad \text{és} \quad Z_4 = Z_5 = Z_6 = 1,$$

akkor a (2.1) becslés alapján

$$Z^* = \lambda_1 \cdot 0 + \lambda_2 \cdot 0 + \lambda_3 \cdot 0 + \lambda_4 \cdot 1 + \lambda_5 \cdot 1 + \lambda_6 \cdot 1 = -0,111$$

adódna. Vagyis, egy esetleg szükségképpen nemnegatív függvény optimális becslése negatív értéknek is adódhat, amely ellentmondana a vizsgált függvény, paraméter természetének. Például negatív telepvastagság, negatív kalóriatartalom nem értelmezhetők.

Az itt felvetett problémán úgy segíthetünk legegyszerűbben, ha megköveteljük, hogy a λ_i súlyok valamennyien nemnegatívak legyenek. Minthogy egy λ_i együttható akkor és csak akkor nemnegatív, ha felírható egy valós szám négyzeteként (vagyis $\lambda_i = a_i^2$), e többszörfeltétel mellett az optimális becslési együtthatók a

$$(2.6) \quad -\gamma_{vv} + 2 \sum_{i=1}^n a_i^2 \gamma_{vi} - \sum_{i=1}^n \sum_{j=1}^n a_i^2 a_j^2 \gamma_{ij} \rightarrow \min$$

$$\sum_{i=1}^n a_i^2 - 1 = 0$$

feltételes szélsőérték probléma megoldásából adódnak. A célfüggvény a becslési varianciával azonos, a feltétel pedig a becslés torzítatlanságát jelenti. A *Lagrange-féle multiplikátor elv* (SZÉP, 1974) felhasználásával az

$$(2.7) \quad L(a_1, \dots, a_n, \mu) = -\gamma_{vv} + 2 \sum_{i=1}^n a_i^2 \gamma_{vi} - \sum_{i=1}^n \sum_{j=1}^n a_i^2 a_j^2 \gamma_{ij} - 2\mu \left(\sum_{i=1}^n a_i^2 - 1 \right)$$

Lagrange-féle függvény feltételek nélküli szélsőértékfeladatát nyerjük. Egyszerű deriválással a következő nemlineáris egyenletek adódnak

$$(2.8) \quad \frac{\partial L}{\partial a_i} = 4a_i \left(\gamma_{vi} - \sum_{j=1}^n a_j^2 \gamma_{ij} - \mu \right) = 0 \quad (i = 1, \dots, n)$$

$$(2.9) \quad \frac{\partial L}{\partial \mu} = -2 \left(\sum_{i=1}^n a_i^2 - 1 \right) = 0.$$

Tekintsük először a (2.8) egyenletet. Ha $a_i = 0$, akkor az egyenlőség nyilvánvalóan fennáll, ha pedig $a_i \neq 0$, akkor

$$(2.10) \quad \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu = \gamma_{vi},$$

vagyis a (2.2) egyenletrendszer i -edik egyenlete teljesül. A (2.9) feltétel pedig az a számok definíciója alapján azonos az eredeti feltétellel. Ez az észrevétel a következőképpen is megfogalmazható. Legyen $\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*$ az optimális nemnegatív súlyok rendszere (azaz a (2.6) feltételes szélsőérték probléma optimális megoldása). Jelölje $\lambda_{i_1}^*, \dots, \lambda_{i_r}^*$ a szigorúan pozitív súlyokat, ekkor ezek megkaphatók a *Kriging módszer* eredeti változatának az alkalmazásával is, ha a megfelelő x_{i_1}, \dots, x_{i_r} mérési pontokat és a Z_{i_1}, \dots, Z_{i_r} függvényértékeket tekintjük csak.

3. A számítási algoritmus

Az előzőekben részletezett matematikai eredmények alapján a következő véges algoritmust javasolhatjuk az optimális nemnegatív súlyok meghatározására:

1. *Lépés.* Generáljuk az $\{1, 2, \dots, n\}$ halmaz összes nem üres $\{i_1, i_2, \dots, i_r\}$ részhalmazát.

2. *Lépés.* Minden egyes $\{i_1, i_2, \dots, i_r\}$ részhalmaz esetére alkalmazzuk a *Kriging módszert* az $x_{i_1}, x_{i_2}, \dots, x_{i_r}$ alappontokra és a $Z_{i_1}, Z_{i_2}, \dots, Z_{i_r}$ függvényértékekre alapozva. Ha az összes súly nemnegatív, akkor számítsuk ki a becslési varianciát a

$$(3.1) \quad V_{i_1 i_2 \dots i_r} = -\gamma_{vv} + 2 \sum_{k=1}^r \lambda_{i_k} \gamma_{vi_k} - \sum_{k=1}^r \sum_{l=1}^r \lambda_{i_k} \lambda_{i_l} \gamma_{i_k i_l}$$

képlet felhasználásával. Ha legalább egy súly negatívnak adódott, akkor generáljuk a következő részhalmazt.

3. *Lépés.* Azt a részhalmazt és megfelelő pozitív λ_k^* súlyokat fogadjuk el optimálisnak, amelyek mellett a legkisebb (3.1) varianciaérték adódott. Az $\{i_1, i_2, \dots, i_r\}$ halmazhoz nem tartozó λ_j^* súlyokat vegyük zérusnak.

Ennek az algoritmusnak az alkalmazását elsősorban az nehezíti, hogy minden egyes $\{i_1, i_2, \dots, i_r\}$ részhalmaz esetén újra végig kell számolnunk a teljes *Kriging módszert*, amely minden egyes esetben $O(r^3)$ számú művelet elvégzését igényli (l. például SZIDAROVSKY, 1974, SZIDAROVSKY és YAKOWITZ, 1978). A lehetséges, nem üres $\{i_1, i_2, \dots, i_r\}$ részhalmazok száma is viszonylag nagy: $2^n - 1$. Például, az $n = 6$ esetben $2^6 - 1 = 63$, és például $n = 10$ esetében $2^{10} - 1 = 1023$.

A továbbiakban azt mutatjuk ki, hogy a lehetséges nem üres részhalmazok szisztematikus generálásával és a *Kriging módszer* egy alkalmas adaptív változatának (l. SZIDAROVSKY és MOLNÁR, 1982) a felhasználásával a fenti algoritmus jelentősen leegyszerűsíthető, a szükséges számítások mennyisége nagyságrendekkel lecsökkenthető. Vizsgálatunkat három lépésre bontjuk a továbbiakban.

A) Új mérési pont bevonása

A (2.2) egyenletrendszer sorainak és oszlopainak alkalmas átrendezésével kimutatható, hogy az $x_{i_1}, x_{i_2}, \dots, x_{i_r}$ alappontokra épülő *Kriging-rendszer* a

$$(3.2) \quad \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & \lambda_{i_1 i_1} & \gamma_{i_1 i_2} & \dots & \gamma_{i_1 i_r} \\ 1 & \lambda_{i_2 i_1} & \gamma_{i_2 i_2} & \dots & \gamma_{i_2 i_r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{i_r i_1} & \gamma_{i_r i_2} & \dots & \gamma_{i_r i_r} \end{pmatrix} \begin{pmatrix} \mu \\ \lambda_{i_1} \\ \lambda_{i_2} \\ \vdots \\ \lambda_{i_r} \end{pmatrix} = \begin{pmatrix} 1 \\ \gamma_{vi_1} \\ \gamma_{vi_2} \\ \vdots \\ \gamma_{vi_r} \end{pmatrix}$$

alakban is felírható. Jelölje G az együttható mátrixot, λ az ismeretlenekből alkotott vektort és γ a jobb oldali vektort, ekkor (3.2) tömören a

$$(3.3) \quad G\lambda = \gamma$$

formában is reprezentálható. Ha az újabb $x_{i_{r+1}}$ mérési pontot is bevonjuk a *Kriging-rendszerbe*, akkor a kibővített lineáris egyenletrendszer a következő:

$$(3.4) \quad \begin{pmatrix} G & g \\ g^T & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \lambda_{i_{r+1}} \end{pmatrix} = \begin{pmatrix} \gamma \\ \gamma_{vi_{r+1}} \end{pmatrix},$$

ahol $\mathbf{g} = (1, \gamma_{i_1 i_{r+1}}, \dots, \gamma_{i_r i_{r+1}})^T$. Kimutatható (I. SZIDAROVSKY (1974, 1981)), hogy ekkor a (3.4) egyenletrendszer együttható mátrixának inverze

$$(3.5) \quad \begin{pmatrix} \mathbf{X} & \mathbf{y} \\ \mathbf{y}^T & s \end{pmatrix},$$

ahol

$$(3.6) \quad s = \frac{-1}{\mathbf{g}^T \mathbf{G}^{-1} \mathbf{g}}, \quad \mathbf{y} = -\mathbf{G}^{-1} \mathbf{g} s, \quad \mathbf{X} = \mathbf{G}^{-1} + \frac{1}{s} \mathbf{y} \mathbf{y}^T,$$

valamint, ha λ_0 jelöli a (3.3) egyenletrendszer megoldását, akkor

$$(3.7) \quad \lambda_{i_{r+1}} = s(\gamma_{v_{i_{r+1}}} - \mathbf{g}^T \lambda_0),$$

$$(3.8) \quad \lambda = \lambda_0 - \mathbf{G}^{-1} \mathbf{g} \lambda_{i_{r+1}}.$$

Továbbá,

$$(3.9) \quad V_{i_1 \dots i_r i_{r+1}} = V_{i_1 \dots i_r} - \frac{(\gamma_{v_{i_{r+1}}} - \mathbf{g}^T \mathbf{G}^{-1} \mathbf{g})^2}{\mathbf{g}^T \mathbf{G}^{-1} \mathbf{g}}.$$

A fenti kifejezéseket a következő sorrendben és változatban célszerű alkalmaznunk a gyakorlatban:

1. Lépés.

$$(3.10) \quad \mathbf{m} = \mathbf{G}^{-1} \mathbf{g};$$

2. Lépés.

$$(3.11) \quad s = \frac{-1}{\mathbf{g}^T \mathbf{m}}; \quad \mathbf{y} = -\mathbf{m} s;$$

$$\mathbf{X} = \mathbf{G}^{-1} + \frac{1}{s} \mathbf{y} \mathbf{y}^T;$$

3. Lépés.

$$(3.12) \quad \lambda_{i_{r+1}} = s(\gamma_{v_{i_{r+1}}} - \mathbf{g}^T \lambda_0),$$

$$\lambda = \lambda_0 - \mathbf{m} \lambda_{i_{r+1}};$$

4. Lépés.

$$(3.13) \quad V_{i_1 \dots i_r i_{r+1}} = V_{i_1 \dots i_r} + \frac{\lambda_{i_{r+1}}^2}{s}.$$

Az itt megadott lépéseket A) algoritmusnak fogjuk a továbbiakban nevezni.

B) Egy mérési pont elhagyása

Tegyük fel ezután, hogy az $x_{i_1}, \dots, x_{i_r}, x_{i_{r+1}}$ pontok közül az $x_{i_{r+1}}$ mérési pontot kihagyjuk a becslési folyamatból. Ekkor feladatunk éppen az előzőekben tárgyalt eset megfordítását jelenti, amikor a (3.5) inverzmátrix, (3.4) megoldása és a $V_{i_1 \dots i_r i_{r+1}}$ variancia ismeretében határozzuk meg a \mathbf{G}^{-1} mátrixot, a λ_0 megoldást és a $V_{i_1 \dots i_r}$ varianciát.

Az előzőek alapján ezeket a számításokat a következő sorrendben célszerű elvégeznünk:

1. Lépés.

$$(3.14) \quad \mathbf{m} = -\frac{1}{s} \mathbf{y};$$

2. Lépés.

$$(3.15) \quad \mathbf{G}^{-1} = \mathbf{X} - \frac{1}{s} \mathbf{y} \mathbf{y}^T;$$

3. Lépés.

$$(3.16) \quad \lambda_0 = \lambda + \mathbf{m} \lambda_{i_r+1};$$

4. Lépés.

$$(3.17) \quad V_{i_1 \dots i_r} = V_{i_1 \dots i_r i_{r+1}} - \frac{\lambda_{i_r+1}^2}{s}.$$

Ezt a számítási sorrendet a továbbiakban B) algoritmusnak fogjuk nevezni.

C) A szükséges részhalmazok rendszeres leírása

Az összes lehetséges, nem üres $\{i_1, \dots, i_r\}$ részhalmazt egy irányított fával ábrázolhatjuk a következőképpen. A fa kezdőpontjának a teljes $N = \{1, 2, \dots, n\}$ halmaz felel meg és minden nem üres $\{i_1, \dots, i_r\}$ részhalmazhoz a fa egy-egy további csúcsa rendelhető. Az $I = \{i_1, \dots, i_r\}$ részhalmaznak megfelelő csúcsból akkor és csak akkor megy irányított él a $J = \{j_1, \dots, j_s\}$ részhalmaznak megfelelő csúcsba, ha az $N - I$ halmazból annak legnagyobb elemének elhagyásával kaphatjuk meg az $N - J$ halmazt. Ezt más szavakkal úgy is kifejezhetjük, hogy az irányított élek kezdőpontjaihoz rendelt halmazokból egy olyan további pont elhagyásával kaphatjuk meg a végpontokhoz rendelt halmazokat, amelyek indexe nagyobb, mint az eddig elhagyott pontok indexei. Az 1. ábrán ezt az irányított fát ábrázoljuk az $n=4$ speciális esetben. Például az $I = \{1, 3, 4\}$ és $J = \{1, 4\}$ csúcsokat köti össze él, hiszen $N - I = \{2\}$ és $N - J = \{2, 3\}$, valamint $N - J$ legnagyobb eleme 3, és annak elhagyásával adódik az $N - I$ halmaz.

Az irányított gráf csúcsainak rendszeres vizsgálata és minden szükséges csúcs esetére a *Kriging-módszer* súlyainak és a becslési varianciának a számítása a következő lépésekből áll:

1. Lépés. Számoljuk végig a *Kriging-módszert* az összes pont bevonása esetére. Ha minden súly nemnegatívnak adódik, akkor az eljárás véget ér. Különben menjünk a következő lépésre.

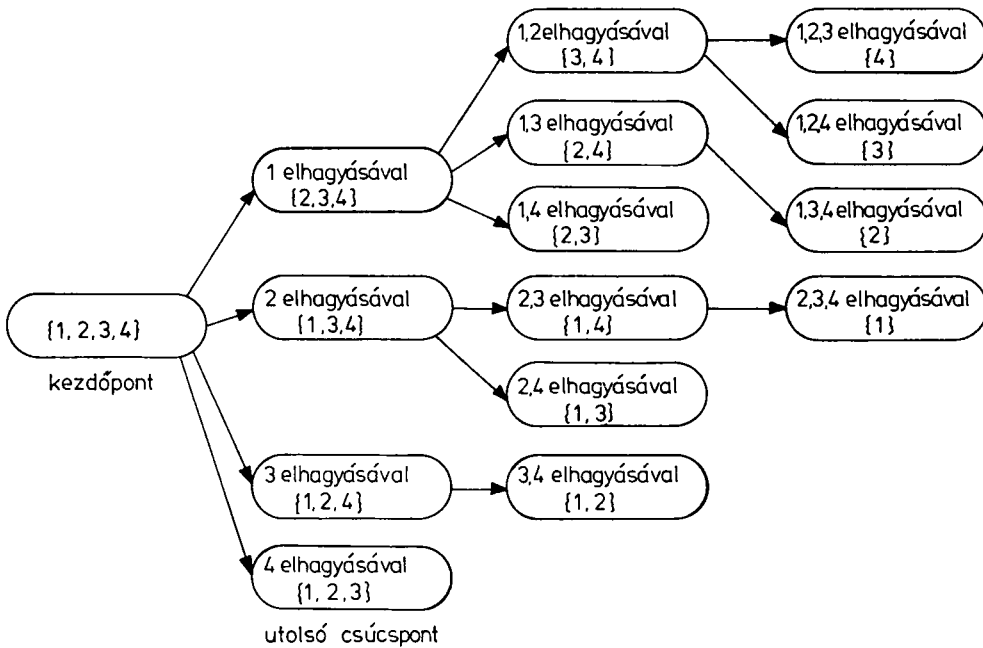
2. Lépés. Legyen $s := 0$, $t := 0$, $M := \infty$, $M_1 := \infty$, $K := 0$.

3. Lépés. Ha az

$$s = n-1, \quad j_s = n, \quad t = n, \quad M_1 > M, \quad K = 1$$

feltételek legalább egyike teljesül, akkor legyen $K := 0$ és menjünk a következő lépésre. Ellenkező esetben pedig menjünk az 5. lépésre.

4. Lépés. Legyen $t := j_s$, $s := s - 1$ és menjünk a 7. lépésre.



1. ábra

Az irányított fa szerkezete $n=4$ esetén

5. Lépés. Legyen $t := t+1$, $s := s+1$, $j_s := t$ és menjünk a következő lépésre.

6. Lépés. Alkalmazzuk a B) algoritmust olyan feltétel mellett, hogy a j_{s+1} pontot hagyjuk el a becslésből. Őrizzük meg az eredményeket a számítógép memóriájában. Ha mindegyik új súly nemnegatívnak adódott, akkor jelölje M_1 az új becslési varianciát. Ha $M_1 < M$, akkor legyen $M := M_1$ és a nyert súlyokat őrizzük meg a memóriában, mint az eddigi legjobb súlyokat. Ha $M_1 \geq M$, akkor az előzőekben nyert legjobb súlyokat hagyjuk változatlanul. Legyen $K := 1$ és menjünk a 8. lépésre.

7. Lépés. Az új Kriging adatokat és eredményeket vegyük elő a számítógép memóriájából, majd menjünk a 3. lépésre.

8. Lépés. Ha $s=1$ és $j_s=n$, akkor az eljárás véget ér. Különben menjünk a 3. lépésre.

Az itt vázolt lépéseket C) algoritmusnak nevezzük a továbbiakban. Az algoritmus lépései a következőképpen magyarázhatók. Az 1. és 2. lépés az irányított gráf egyetlen kezdőpontjához tartoznak. Itt s jelenti az elhagyott pontok számát, ez nyilvánvalóan 0 a kezdőpont esetében. A t változó adja meg azt, hogy melyik volt a legnagyobb indexű pont, amelyet ebből a csúcsból kiindulva a korábbi lépésekben elhagytunk. M_1 jelenti az aktuális becslési varianciát és M jelöli az aktuális lépésig talált legjobb varianciát. Ha $K=1$, akkor az éppen megvizsgált csúcspontban, minden súly nemnegatív, az ellenkező esetben pedig $K=0$. A 3. lépésben azt vizsgáljuk meg, hogy az

éppen vizsgált csúcspontból tovább tudunk-e menni egy következő csúcspontra. Ha $s=n-1$, akkor egy híján az összes mérési pontot elhagytuk, így további pont elhagyásával nem kaphatunk nem üres részhalmazt. Ha $j_s=n$, akkor az I halmazból már elhagytuk a lehetséges legnagyobb indexű pontot, így még nagyobb indexű újabb pont már nem hagyható el. Ha $t=n$, akkor az ebből a csúcsból kiinduló élek végpontjait már mind sorra vettük az előzőekben, így ebből a csúcsból nem tudunk továbbmenni előre. Ha $M_1 > M$, akkor az éppen vizsgált csúcshoz nemnegatív súlyok tartoznak és a megfelelő becslési variancia rosszabb, mint az előzőekben talált legjobb érték. Ilyenkor sem célszerű e csúcsból továbbhaladni előre, mert további mérési pontok elhagyásával nem javulhat a becslési variancia értéke, így semmilyen további csúcspont sem lehet optimális. A $K=1$ esetben az éppen megvizsgált csúcshoz tartozó összes súly nemnegatív. Ilyenkor sem kell továbbmennünk, mert további mérési pontok elhagyásával a becslési variancia már nem javulhat. A 4. lépés az illető csúcspontból az oda befutó el mentén való visszalépést jelenti. Az 5. lépés az illető csúcsból való előrelépést jelenti, amikor olyan újabb mérési pontot hagyunk el, amely az előzőekben már elhagyottak után következik. A 6. és 7. lépés az új csúcsponthoz tartozó részhalmaznak megfelelő *Kriging-módszert* hajtja végre, a 6. lépésben a B) algoritmust kell felhasználnunk, a 7. lépésben (minthogy visszafelé lépés történt és ezért ezt az újabb csúcsot már megvizsgáltuk az egyik korábbi lépés során) pedig a memóriából kell csak előszedni a már korábban kiszámolt mennyiségeket. Ha a számítógép memóriája korlátozott oly módon, hogy nem tudjuk a korábbi eredményeket tárolni benne, akkor a 7. lépést úgy kell módosítanunk, hogy az A) algoritmus felhasználásával újra számítsuk ki a *Kriging-módszer* eredményeit. A 8. lépésben azt vizsgáljuk meg, hogy a vizsgálat során elértük-e az utolsó pontot. Ez a pont az ábrán az $\{1, 2, 3\}$ részhalmaznak felel meg. Innen továbbhaladni már nem lehet, visszafelé lépés után is csak olyan további csúcspontokban haladhatunk, amelyeket már megvizsgáltunk a korábbi lépések során. Például, az 1. ábrán bemutatott fa esetében az egymásután megvizsgált csúcsok a következők:

$$\begin{aligned} &\{1, 2, 3, 4\} \rightarrow \{2, 3, 4\} \rightarrow \{3, 4\} \rightarrow \{4\} \leftarrow \{3, 4\} \rightarrow \{3\} \leftarrow \{3, 4\} \leftarrow \{2, 3, 4\} \rightarrow \\ &\rightarrow \{2, 4\} \rightarrow \{2\} \leftarrow \{2, 4\} \leftarrow \{2, 3, 4\} \rightarrow \{2, 3\} \leftarrow \{2, 3, 4\} \leftarrow \{1, 2, 3, 4\} \rightarrow \{1, 3, 4\} \rightarrow \\ &\rightarrow \{1, 4\} \rightarrow \{1\} \leftarrow \{1, 4\} \leftarrow \{1, 3, 4\} \rightarrow \{1, 3\} \leftarrow \{1, 3, 4\} \leftarrow \{1, 2, 3, 4\} \rightarrow \{1, 2, 4\} \rightarrow \\ &\rightarrow \{1, 2\} \leftarrow \{1, 2, 4\} \leftarrow \{1, 2, 3, 4\} \rightarrow \{1, 2, 3\}. \end{aligned}$$

Itt természetesen feltettük, hogy az összes csúcs megvizsgálásra kerül (azaz a becslési variancia viszonylag nagy értéke miatt nem marad ki teljes lánc a vizsgálatból). Itt \rightarrow az előre haladást, \leftarrow pedig a visszafelé haladást jelenti. Mint láthatjuk, ugyanaz a csúcs többször is szerepel, a hozzá tartozó eredmények újra történő számítása csak akkor célszerű, ha a számítógép memóriájában ezek tárolása nem lehetséges.

A C) algoritmus konkrét alkalmazását mutatjuk be a következőkben.

3.1. PÉLDA. A 2.1. példában láttuk, hogy az $x_1=(0,20)$, $x_2=(20,20)$, $x_3=(0,60)$, $x_4=(0,70)$, $x_5=(0,0)$, $x_6=(0,-40)$ pontokból kiindulva a $V=(10,35)$ pontbeli függvényérték számításakor negatív súlyok is fellépnek, ha a

$$\gamma(h) = \begin{cases} \frac{3}{2} \frac{h}{80} - \frac{1}{2} \frac{h^3}{80^3}, & \text{ha } 0 \leq h \leq 80 \\ 1, & \text{ha } h \geq 80 \end{cases}$$

variogramot feltételezzük. Tehát az optimális nemnegatív súlyok az $\{1, 2, 3, 4, 5, 6\}$ halmaz valamilyen nem üres részhalmazához tartoznak. A C) algoritmus felhasználásával meghatároztuk az optimális nemnegatív súlyokat. A részeredményeket az 1. táblázat tartalmazza. A vizsgálatba bevont részhalmazokat a táblázat 1. oszlopa tartalmazza a bevonás sorrendjében. A második oszlop adja meg a nemnegatív súlyokat. Amennyiben egy részhalmaz esetén negatív súly is adódott, akkor a hozzátartozó súlyokat nem adjuk meg. Meg kell jegyeznünk, hogy a részhalmazokba nem tartozó mérési pontokhoz zérus súlyok tartoznak. A táblázatból közvetlenül leolvasható, hogy a nemnegatív súlyok melletti legkisebb becslési variancia értéke 0,364, amelyet az $\{1, 2, 3\}$ részhalmaz esetén kapunk. Tehát az optimális nemnegatív súlyok a következők:

1. TÁBLÁZAT

A számítás részeredményei

Részhalmaz	Együtthatók	Becslési variancia
1; 2; 3; 4; 5; 6	—	0,358
2; 3; 4; 5; 6	—	0,420
3; 4; 5; 6	—	0,651
4; 5; 6	—	0,782
5; 6	0,762; 0,238	1,193
4; 6	0,680; 0,320	1,066
4; 5	0,500; 0,500	0,782
3; 5; 6	—	0,651
3; 6	0,755; 0,245	0,851
3; 5	0,582; 0,418	0,652
3; 4; 6	—	0,851
3; 4	0,901; 0,099	0,968
3; 4; 5	—	0,651
2; 4; 5; 6	—	0,478
2; 5; 6	0,775; 0,132; 0,093	0,612
2; 4; 6	0,670; 0,318; 0,012	0,485
2; 4; 5	0,610; 0,304; 0,086	0,479
2; 3; 5; 6	—	0,420
2; 3; 6	0,602; 0,397; 0,001	0,426
2; 3; 5	0,547; 0,381; 0,072	0,422
2; 3; 4; 6	—	0,426
2; 3; 4; 5	—	0,421
1; 3; 4; 5; 6	—	0,457
1; 2; 4; 5; 6	—	0,390
1; 2; 5; 6	—	0,467
1; 2; 4; 6	—	0,397
1; 2; 4	0,369; 0,396; 0,235	0,399
1; 2; 3; 5; 6	—	0,358
1; 2; 3; 6	—	0,362
1; 2; 3	0,321; 0,373; 0,306	0,364
1; 2; 3; 5	—	0,358
1; 2; 3; 4; 6	—	0,361
1; 2; 3; 4	—	0,364
1; 2; 3; 4; 5	—	0,357

$$\lambda_1 = 0,321; \quad \lambda_2 = 0,373; \quad \lambda_3 = 0,306; \quad \lambda_4 = \lambda_5 = \lambda_6 = 0.$$

Ha a 2.1. példában bemutatott

$$Z_1 = Z_2 = Z_3 = 0, \quad Z_4 = Z_5 = Z_6 = 1$$

függvényértékekből indulunk ki, akkor a

$$Z^* = \sum_{i=1}^6 \lambda_i Z_i = 0$$

becslés adódik. Ha a függvényértékeket a

$$Z_1 = Z_2 = Z_3 = 1, \quad Z_4 = Z_5 = Z_6 = 4$$

szabállyal adjuk meg, akkor ezzel a módszerrel a $Z^*=1$ becslés adódik. Ha negatív súlyokat is megengedünk, akkor a klasszikus *Kriging módszerrel* számolva a $Z^*=0,665$ eredmény adódik. Tehát nemcsak a becslési variancia, hanem a becslt függvényérték szempontjából is különbséget mutat a két módszer.

4. A módszer egy módosítása

Az előzőekben nemnegatív súlyokkal kívántuk elérni azt, hogy a Z^* becslés a legnagyobb és legkisebb mért függvényérték közé essék. A becslésre adandó alsó, ill. és felső korlátok más módon is figyelembe vehetők. Ekkor is a (2.1) lineáris kifejezéssel becsljük a Z függvény V tartománybeli átlagértékét, ahol az ismeretlen λ_j ($1 \leq j \leq n$) együtthatókat a következő három kritérium alapján határozzuk meg:

a) A Z^* becslés torzítatlan kell legyen, amely ekvivalens a

$$(4.1) \quad \sum_{i=1}^n \lambda_i - 1 = 0$$

feltétellel (DAVID, 1977).

b) A Z^* becslés korlátos kell legyen, vagyis

$$(4.2) \quad Z_{\min} \leq Z^* \leq Z_{\max},$$

ahol Z_{\min} és Z_{\max} adott értékek. Például, a $Z^* \geq 0$ nemnegativitási feltétel esetén

$$(4.3) \quad Z_{\min} = 0 \quad \text{és} \quad Z_{\max} = \infty.$$

c) A Z^* becslés bizonytalanságát mutató variancia minimális kell legyen, azaz

$$(4.4) \quad \text{Var}[Z^*] \rightarrow \min.$$

A (2.4) célfüggvény (4.1) és (4.2) feltételek melletti optimalizálására a *Lagrange-féle multiplikátor módszert* (SZÉP, 1974) alkalmazhatjuk, amihez az összes feltételt egyenlőség alakra kell hoznunk. A (4.1) feltétel eleve egyenlőség alakjában adott, a (4.2) feltétel pedig nyilvánvalóan ekvivalens a

$$(4.5) \quad \sum_{i=1}^n \lambda_i Z_i + \alpha_1^2 - Z_{\max} = 0$$

$$\sum_{i=1}^n \lambda_i Z_i - \alpha_2^2 - Z_{\min} = 0$$

egyenlőség-párral, ahol α_1, α_2 ismeretlen konstansok. Esetünkben a *Lagrange-függvény* a következő alakú:

$$(4.6) \quad L(\lambda_1, \dots, \lambda_n, \alpha_1, \alpha_2, \mu_1, \mu_2, \mu_3) = -\gamma_{vv} + 2 \sum_{i=1}^n \lambda_i \gamma_{vi} - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma_{ij} - \\ - 2\mu_1 \left(\sum_{i=1}^n \lambda_i Z_i + \alpha_1^2 - Z_{\max} \right) - 2\mu_2 \left(\sum_{i=1}^n \lambda_i Z_i - \alpha_2^2 - Z_{\min} \right) - 2\mu_3 \left(\sum_{i=1}^n \lambda_i - 1 \right),$$

amely parciális deriváltjai az alábbiak:

$$(4.7) \quad \frac{\partial L}{\partial \lambda_i} = 2\gamma_{vi} - 2 \sum_{j=1}^n \lambda_j \gamma_{ij} - 2\mu_1 Z_i - 2\mu_2 Z_i - 2\mu_3 \quad (\forall i)$$

$$(4.8) \quad \frac{\partial L}{\partial \alpha_1} = -2\mu_1 \alpha_1$$

$$(4.9) \quad \frac{\partial L}{\partial \alpha_2} = 2\mu_2 \alpha_2$$

$$(4.10) \quad \frac{\partial L}{\partial \mu_1} = \sum_{i=1}^n \lambda_i Z_i + \alpha_1^2 - Z_{\max}$$

$$(4.11) \quad \frac{\partial L}{\partial \mu_2} = \sum_{i=1}^n \lambda_i Z_i - \alpha_2^2 - Z_{\min}$$

$$(4.12) \quad \frac{\partial L}{\partial \mu_3} = \sum_{i=1}^n \lambda_i - 1.$$

A parciális deriváltak zérus volta a

$$(4.13) \quad \begin{aligned} \gamma_{vi} - \sum_{j=1}^n \lambda_j \gamma_{ij} - \mu_1 Z_i - \mu_2 Z_i - \mu_3 &= 0 \quad (\forall i) \\ \mu_1 \alpha_1 &= 0 \\ \mu_2 \alpha_2 &= 0 \\ \sum_{j=1}^n \lambda_j Z_j + \alpha_1^2 - Z_{\max} &= 0 \\ \sum_{j=1}^n \lambda_j Z_j - \alpha_2^2 - Z_{\min} &= 0 \\ \sum_{j=1}^n \lambda_j - 1 &= 0 \end{aligned}$$

nemlineáris egyenletrendszer megoldására vezet, amely megoldására rutinszerűen alkalmazható eljárások állnak rendelkezésre. (SZIDAROVSKY, 1974, SZIDAROVSKY és YAKOWITZ, 1978). A következőkben a (4.13) egyenletrendszer megoldását lineáris egyenletrendszer numerikus megoldására vezetjük vissza.

A (4.13) egyenletrendszer második és harmadik egyenlete alapján a megoldás ki kell elégítse a következő négy eset valamelyikét:

A) $\mu_1 = \mu_2 = 0;$

B) $\mu_1 = \alpha_2 = 0;$

C) $\mu_2 = \alpha_1 = 0;$

D) $\alpha_1 = \alpha_2 = 0.$

Az A) esetben a (4.13) egyenletrendszer a következővel ekvivalens:

$$(4.14) \quad \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu_3 = \gamma_{vi}; \quad (\forall i) \quad \sum_{j=1}^n \lambda_j = 1; \quad \sum_{j=1}^n \lambda_j Z_j \leq Z_{\max}; \quad \sum_{j=1}^n \lambda_j Z_j \geq Z_{\min};$$

Az első két egyenlet a klasszikus *Kriging-módszer* egyenletrendszerével azonos, így ebből a λ_i és μ_3 ismeretlenek közvetlenül megkaphatók. Nyilvánvaló, hogy ez az egyenletrendszer tömören a

$$(4.15) \quad \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_3 \end{pmatrix} = \begin{pmatrix} \gamma_{v1} \\ \vdots \\ \gamma_{vn} \\ 1 \end{pmatrix}$$

alakban írható fel. Ha a megoldás kielégíti a két utolsó egyenlőtlenséget, akkor feladatunk megoldását is jelenti, ha nem, akkor ez az eset nem teljesül.

A B) esetben a (4.13) egyenletrendszer

$$(4.16) \quad \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu_2 Z_i + \mu_3 = \gamma_{vi} \quad (\forall i); \quad \sum_{j=1}^n \lambda_j Z_j = Z_{\min};$$

$$\sum_{j=1}^n \lambda_j = 1 \quad \sum_{j=1}^n \lambda_j Z_j \leq Z_{\max}$$

alakú. Az utolsó feltétel nyugodtan elhagyható, mert a másodiknak egyszerű következménye. Az első három egyenlet lineáris, $n+2$ összefüggést ad $n+2$ ismeretlenre, valamint tömören a

$$(4.17) \quad \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 & Z_1 \\ \vdots & & \vdots & \vdots & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 & Z_n \\ 1 & \dots & 1 & 0 & 0 \\ Z_1 & \dots & Z_n & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_3 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \gamma_{v1} \\ \vdots \\ \gamma_{vn} \\ 1 \\ Z_{\min} \end{pmatrix}$$

mátrix alakban írható fel. Vegyük észre, hogy az együttható mátrix szimmetrikus, valamint $Z^* = Z_{\min}$.

A C) esetben az előzőhöz hasonlóan a

$$(4.18) \quad \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu_1 Z_i + \mu_3 = \gamma_{vi} \quad (\forall i); \quad \sum_{j=1}^n \lambda_j Z_j = Z_{\max};$$

$$\sum_{j=1}^n \lambda_j = 1; \quad \sum_{j=1}^n \lambda_j Z_j \cong Z_{\min}$$

relációkat kapjuk, ahol az utolsó egyenlőtlenség most is egyszerű következménye a másodiknak, így egyszerűen elhagyható. A (4.18) egyenletrendszer a (4.17)-hez hasonló

$$(4.19) \quad \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 & Z_1 \\ \vdots & & \vdots & \vdots & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 & Z_n \\ 1 & \dots & 1 & 0 & 0 \\ Z_1 & \dots & Z_n & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_3 \\ \mu_1 \end{pmatrix} = \begin{pmatrix} \lambda_{v1} \\ \vdots \\ \gamma_{vn} \\ 1 \\ Z_{\max} \end{pmatrix}$$

tömör alakban is felírható. Vegyük észre, hogy a (4.17) és (4.19) egyenletrendszerek együtttható mátrixa azonos, valamint ez esetben $Z^* = Z_{\max}$.

A D) eset soha sem teljesülhet, hiszen ekkor a (4.13) negyedik és ötödik egyenlete alapján

$$(4.20) \quad \sum_{j=1}^n \lambda_j Z_j - Z_{\max} = 0$$

$$\sum_{j=1}^n \lambda_j Z_j - Z_{\min} = 0,$$

amely nyilvánvaló ellentmondás $Z_{\min} < Z_{\max}$ esetén.

Az elmondottakat a következő algoritmusban foglalhatjuk össze:

1. *Lépés.* A klasszikus *Kriging-módszer* alkalmazásával becsüljük meg a Z_v értéket. Ha a becslés kielégíti a (4.2) feltételeket, akkor az eljárás véger ér, a becslést elfogadjuk és a becslés bizonytalanságát a klasszikus *Kriging-módszer* alkalmazásakor nyert becslési varianciával jellemezzük. Ha a becslés nem elégíti ki a (4.2) feltételt, akkor menjünk a következő lépésre.

2. *Lépés.* A klasszikus *Kriging-módszerrel* adódó becslést toljuk el a $[Z_{\min}, Z_{\max}]$ intervallum határára, ekkor a becslési varianciát a (2.4) formula adja meg, ahol a λ_i együttthatókat a megfelelő (4.17) vagy (4.19) egyenletrendszer megoldásából kapjuk meg.

Vegyük észre, hogy a (4.17) és (4.19) egyenletrendszer úgy kapható meg a klasszikus *Kriging-módszer* (2.2) egyenletrendszeréből, hogy annak együttthatómátrixához egy újabb sort és oszlopot adunk, valamint a jobb oldali vektort is kiegészítjük egy újabb elemmel. Ekkor pedig a (4.15) rendszer megoldásából a (4.17) és (4.19) egyenletrendszer megoldása könnyen kiadódik, ha az előzőekben bemutatott eljárást itt is alkalmazzuk. Ez az észrevétel lényegesen lecsökkenti a módszer alkalmazásának számításiigényét.

4.1. *Példa.* Alkalmazzuk most az imént részletezett eljárást a 3.1. példában bemutatott feladat esetére.

Ekkor az 1. lépés eredményeként a

$$\lambda_1 = 0,382; \lambda_2 = 0,403; \lambda_3 = 0,326;$$

$$\lambda_4 = -0,013; \lambda_5 = -0,085; \lambda_6 = -0,013$$

együtthatók adódnak. Ha a függvényértékeket a

$$Z_1 = Z_2 = Z_3 = 1, \quad Z_4 = Z_5 = Z_6 = 4$$

szabály szerint adjuk meg, és a Z^* becslésre csak a nemnegativitási feltételt tesszük (vagyis $Z_{\min}=0, Z_{\max}=\infty$), ekkor a klasszikus *Kriging-módszer* alkalmazásával adódó

$$Z^* = 0,665$$

becslés eleget tesz a feltételeknek, és a becslési variancia értéke:

$$\text{Var}[Z^*] = 0,358.$$

Adjuk meg ezután a függvényértékeket a

$$Z_1 = Z_2 = Z_3 = 0, \quad Z_4 = Z_5 = Z_6 = 4$$

szabállyal. Ekkor a klasszikus *Kriging-módszer* alkalmazása (azaz az algoritmus 1. lépése) negatív függvényértékre vezet, hiszen

$$\sum_{j=1}^6 \lambda_j Z_j = -0,444.$$

Ha a Z^* becslésre ismét megköveteljük a nemnegativitást, akkor a 2. lépés alapján a

$$Z^* = 0$$

becslés adódik. A (4.17) egyenletrendszer megoldásából ez esetben pedig a

$$\text{Var}[Z^*] = 0,364$$

hibavariancia értéket nyerjük. Tehát a két módszer nemcsak a becsült függvényérték szempontjából mutat eltérést, hanem a becslés bizonytalanságát illetően is. Esetünkben a klasszikus *Kriging-módszer* elfogadhatatlan átlagértéket adott nagyobb biztonsággal, mint az itt bevezetett módszer elfogadható (bár triviális) átlagértéket adott kisebb biztonsággal.

IRODALOM

- [1] DAVID, M., *Geostatistical Ore Reserve Estimation* (Elsevier Scientific Publ. Company, New York, 1977).
- [2] DELHOMME, J. P., "Kriging in the hydrosiences", *Advances in Water Resources*, **1** (1978) 251—266.
- [3] GAMBOLATI, G. and VOLPI, G., "Groundwater contour mapping in Venice by stochastic interpolation, I. Theory", *Water Resources Research* **15** (1979) 281—290.
- [4] JOURNEL, A. G. and HUIJBREGTS, CH. J., *Mining Geostatistics* (Academic Press, New York, 1978).
- [5] KAPOLYI, L., *Ásványi eredetű természeti erőforrások rendszer- és függvényszemlélete* (Akadémiai Kiadó, Budapest, 1981).

- [6] KRIGE, D. G., "A statistical approach to some basic mine valuation problems on the witwatersrand", *J. Chem. Metall. Min. Soc. S. Afr.* **52** (1951) 119—139.
- [7] MATHERON, G., "Theorie lognormale de l'échantillonnage systematique des gisements", *Ann. Mines* (1957) 566—584.
- [8] MATHERON, G., "Principles of geostatistics", *Econ. Geol.* **58** (1963) 1246—1266.
- [9] MATHERON, G., *Le Krigeage Universel* (Les Cahiers du Centre de Morphologie Mathematique, Fasc. No. 1., ENSMP, Paris, 1969), 82.
- [10] MOLNÁR, S., SZIDAROVSKY, F. és YAKOWITZ, S., „A Kriging módszer néhány fontos tulajdonsága”, *Alkalmazott Matematikai Lapok* **8** (1982).
- [11] SCHAAP, W. and ST. GEORGE, J. D., "On weights of linear Kriging estimator", *Trans. Inst. Min. Metall (Sect. A: Min. Industry)* **90** (1981) A25—A27.
- [12] SZÉP, J., *Analízis* (Közgazdasági és Jogi Könyvkiadó, Budapest, 1974).
- [13] SZIDAROVSKY, F., *Bevezetés a numerikus módszerekbe* (Közgazdasági és Jogi Könyvkiadó, Budapest, 1974).
- [14] SZIDAROVSKY, F. and YAKOWITZ, S., *Principles and Procedures of Numerical Analysis* (Plenum Publishing Company, New York, 1978).
- [15] SZIDAROVSKY, F., „Az univerzális Kriging módszer néhány fontos tulajdonsága”, *KBFI kutatási jelentés*, 1981 augusztus.
- [16] SZIDAROVSKY, F. és MOLNÁR, S., „Néhány megjegyzés az adaptív Kriging módszerhez”, megjelenés alatt, (1982).

(Beérkezett: 1982. augusztus 10.)

(Átdolgozás után beérkezett: 1983. március 25.)

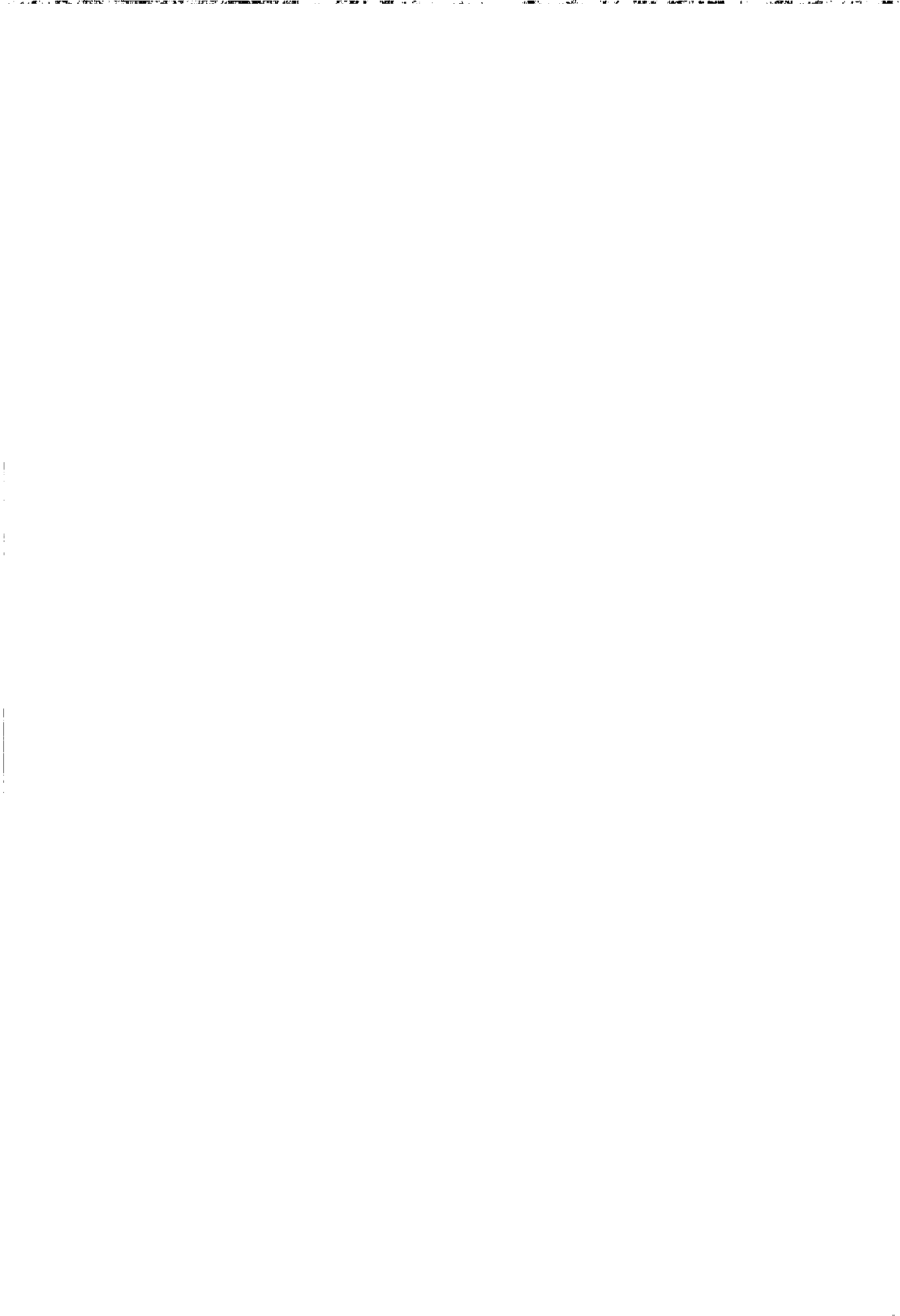
MOLNÁR SÁNDOR
KÖZPONTI BÁNYÁSZATI FEJLESZTÉSI INTÉZET
1037 BUDAPEST, MIKOVINY S. U. 2—4.

SZIDAROVSKY FERENC
KERTÉSZETI EGYETEM MATEMATIKAI ÉS SZÁMÍTÁSTECHNIKAI TANSZÉK
1118 BUDAPEST, VILLÁNYI UT 29—35.

ON THE KRIGING METHOD WITH NONNEGATIV WEIGHTS AND BOUNDS

S. MOLNÁR and F. SZIDAROVSKY

The Kriging method has an important role in applied sciences. It was shown by Schaap and St. George that when applying the method the weights are not necessarily nonnegative. In this paper we improve their method and show that if we accept only nonnegative weights then we get a nonlinear equation system. We give a finite algorithm for the solution of this system.



VELESZÜLETETT RENDELLENESSÉGEK FÜGGETLENSÉGÉNEK VIZSGÁLATA

TELEGDI LÁSZLÓ

Budapest

A dolgozat a Magyarországon 1970—76 között született és a *Veleszületett Rendellenességek Országos Nyilvántartásában* szereplő gyermekek alapján a többszörös veleszületett rendellenességek statisztikai elemzésével foglalkozik. Ezek méretének eloszlását jellemző mennyiségek becslésével vizsgálja meg az egyes veleszületett rendellenességek függetlenségét.

1. A veleszületett rendellenesség definíciója

A fogamzást követően a születésig strukturális, működési és/vagy biokémiai fogyatékoságok alakulhatnak ki az embrióban vagy a magzatban. Egy ilyen, a születés pillanatában létező fogyatékoságot *veleszületett anomáliának* (*congenital anomaly*) nevezünk. Leggyakoribb fajtája a *veleszületett abnormitás* (*congenital abnormality*), amely a hibás morfogenezis következménye, a születés pillanatában létező strukturális fogyatékoság. A veleszületett abnormitás legsúlyosabb — és egyben leggyakoribb — válfaja a *veleszületett malformáció* (*congenital malformation*). A magyar nyelvű szakirodalomban a veleszületett anomália, abnormitás és malformáció kifejezések nem terjedtek el, hanem többnyire mindhármójuk helyett a *veleszületett rendellenesség* kifejezést használják ([8]-ban például a veleszületett malformáció kifejezés helyett). A továbbiakban a veleszületett rendellenesség kifejezés azt a veleszületett abnormitásnál valamivel szélesebb, a veleszületett anomáliánál viszont lényegesen szűkebb kategóriát fogja jelenteni, amelyik hazánkban bejelentésköteles. Ezt a fajta veleszületett rendellenességet a *congenital anomaly* és *congenital abnormality* kifejezések közös kezdőbetűivel CA-nak fogjuk rövidíteni.

Magyarországon 1962-ben indították be (másodiknak a világon) a CA-k folyamatos és kötelező bejelentését, valamint a *Veleszületett Rendellenességek Országos Nyilvántartását* (VRONY). A CA-k osztályozása a VRONY-ban annak 1970-es megújítása óta a *Betegségek Nemzetközi Osztályozásán* (jelenleg ennek 10. Módosításán) alapul, bizonyos változtatásokkal.

2. Többszörös veleszületett rendellenességek

A CA-kon belül különös fontossággal bírnak a *többszörös veleszületett rendellenességek* (*multiple CA*, MCA). Definíciójuknak három kritériuma van.

I. *A CA-k száma.* Korábban az MCA definíciója három vagy több CA együttese volt. Jelen dolgozatban [2] — és néhány más dolgozat, pl. [1] — javaslatára alapján az MCA két vagy több különböző CA együttes előfordulását fogja jelenteni.

II. *A különböző CA-k egymáshoz való viszonya.* Újabban az izolált és a többszörös CA-kat kóreredetük alapján különböztetik meg. Izolált az a CA, amelyik a morfogenezis egyetlen helyhez kötött hibájára vezethető vissza. Az izolált CA-nak három típusa van. Az egyszerű CA olyan strukturális fogyatékoság egyedi megnyilvánulása, amelyik a morfogenezis egyetlen helyhez kötött hibájának a következménye. A komplex CA olyan rendellenességek együttese egy szervben vagy szervrendszerben, amelyek mindegyike a morfogenezis ugyanazon egy helyhez kötött hibájának a következménye. A CA másodlagos CA-kkal (vagy anomalad) a korai morfogenezis egyetlen helyhez kötött hibájának a következménye a későbbi morfogenezis során ebből származó másodlagos rendellenességekkel együtt. — A többszörös CA olyan fejlődési zavar, amely a morfogenezis két vagy több, különböző helyekhez kötött hibájának a következménye. Ennek is három típusa van. A szindróma olyan, feltehetően közös kóreredetű CA-k felismert együttese, amelyek az adott időpontban nem magyarázhatók úgy, mint a morfogenezis ugyanazon egy helyhez kötött hibájának a következményei. A CA-asszociáció CA-k olyan felismert együttese, amelyet az adott időpontban nem tekintenek szindrómának. Végül véletlen kombinációról két vagy több különböző CA véletlen együttes előfordulása esetén beszélünk.

III. *A CA-k jellege.* A CA-k bejelentése során az ún. *minor* CA-kat, amelyek komoly orvosi és kozmetikai következmény nélküli szokatlan morfológiai jellegzetességek, általában nem jelentik be, hanem csak esetlegesen említik. Így ezeket nem lehet figyelembe venni a nyilvántartott MCA-k értékelésénél, csak az ún. *major* CA-kat.

Összefoglalva: definíciója szerint az MCA

- két vagy több
- különböző (vagyis a morfogenezis különböző helyekhez kötött hibáinak következményeként kialakuló)
- major

CA együttes előfordulása ugyanannál a személynél.

3. A veleszületett rendellenességek osztályozása

A VRONY-ban pontosan és részletesen szereplő MCA-k értékeléséhez a major CA-k három kategóriába és a kategóriákon belül csoportokba lettek besorolva ([3], [7]). A kategóriák a hozzájuk tartozó CA-k gyakoriságában, diagnosztizálhatóságában és fontosságában különböznek egymástól. (Az egyes CA-csoportokat két-két nagybetűvel jelöljük.) Az első kategória tartalmazza a *nyilvánvaló* CA-csoportokat (1. táblázat), számuk 11. Az ide tartozó CA-k általában szabad szemmel láthatóak, vagy ha nem, meglétük akkor is nyilvánvaló rögtön a születés után. A második kategória tartalmazza a *fontos* CA-csoportokat (2. táblázat), számuk 9. Az ide tartozó CA-k közepes biztonsággal diagnosztizálhatóak. A harmadik kategória tartalmazza a *kétes* CA-csoportokat (3. táblázat), számuk 20. Az ide tartozó CA-k diagnózisa főleg az orvos gyakorlatától és tudásától függ. (Egy negyedik kategória tartalmazza a *minor* CA-kat. Ezeket azonban, mint fentebb említettük, nem lehet figyelembe venni

1. TÁBLÁZAT

AN	= koponyahiány
EN	= agysérv
SB	= nyitott gerinc
CL	= nyúlajak
CP	= farkastorok
LR	= végtagredukció
PY	= többujjúság
SY	= összenőtt ujjak
EX	= nyitott has
OA	= nyelőcsőelzáródás vagy -szűkület
AA	= végbélelzáródás vagy -szűkület

2. TÁBLÁZAT

AM	= kisszeműség vagy szemnélküliség
CT	= szürkehályog
HS	= húgycsőhasadék
EG	= a külső nemi szervek egyéb rendellenességei
HD	= szívrendellenességek
DI	= rekeszizomsérv
RA	= vesehiány
CK	= veseciszta
AI	= vékonybélelzáródás

3. TÁBLÁZAT

MC	= kisfejűség
HY	= vízfejűség
ON	= az agy egyéb reduktív rendellenességei, további idegrendszeri rendellenességek
FN	= az arc, az orr, a nyelv, a nyak és a koponya további rendellenességei
EY	= további szemrendellenességek
EA	= fülrendellenességek
TC	= ferde nyak
BR	= nyaki ciszták, sipolyok és fül előtti függelék
CD	= csípőficam
CF	= dongaláb
OL	= további végtagrendellenességek
RS	= a légzőszervek rendellenességei
PS	= a gyomorkimenet szűkülete
OD	= további zsigeri rendellenességek
UT	= le nem szállt here
OU	= további húgy-ivarrendszeri rendellenességek
IH	= lágyéksérv
MS	= további mozgásszervi rendellenességek
EO	= az endokrin szervek rendellenességei
ST	= bőr-, haj- és körömrendellenességek, daganatok és további rendellenességek

az MCA-k értékelésénél.) — A CA-k mindössze $n=40$ csoportba történő besorolása (az egyszerűség kedvéért a továbbiakban ezeket a CA-csoportokat fogjuk CA-knak nevezni) természetesen bizonyos információvesztést jelent, és csökkenti az azonosítás pontosságát, ugyanakkor azonban megkönnyíti az MCA-k statisztikai leírását.

4. Az adatmező és tárolása

Magyarországon 1970–76 között 1 188 529 gyermek született. Közülük 38 023 szerepel a VRONY-ban (major vagy minor) CA-s gyermekként. Ezek közül 4959-nek (13,0 százalék) több CA-ja volt. Az ilyen esetek 31,7 százalékában (1573 gyermek) ismert szindróma megállapítása miatt, 3,6 százalékában (180 gyermek) egyéb okból a konkrét CA-k nincsenek részletezve; ezt az 1753 gyermeket ki kellett zárni a statisztikai értékelésből. A CA-s gyermekek közül 5420, ill. 444 volt olyan, akinek csak minor, ill. egy major és egy vagy több minor CA-ja volt. Őket úgy vettük figyelembe, mint akiknek nincs CA-ja, ill. egy CA-ja van. Így végül 30 850 (major) CA-s és ezen belül 2762 MCA-s gyermek maradt $M=1\ 186\ 776$ gyermek közül a többszörös veleszületett rendellenességek statisztikai értékeléséhez. (Az időkorlátozásnak, annak, hogy az utolsó vizsgált év 1976, az a magyarázata, hogy a munka 1977 végén kezdődött az akkor meglévő, 1970–76-os adatokkal, és az adatmezőt később nem lett volna célszerű megváltoztatni. Az 1973 óta működő *Veleszületett Rendellenességek Országos Monitorának* havonkénti értékelése ezt alátámasztotta.) A szükséges, SIMULA nyelvű programok a Magyar Tudományos Akadémia CDC 3300 típusú számítógépén futottak.

Jelölje A_1, A_2, \dots, A_n az egyes CA-kat. Tetszőleges

$$G_\mu = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}, \quad 1 \leq k \leq n, \quad 1 \leq i_1 < i_2 < \dots < i_k \leq n,$$

esetén [ez $(2^n - 1)$ számú G_μ -t jelent, ezért feltehető, hogy $1 \leq \mu \leq 2^n - 1$] jelölje

$$O(G_\mu) = O(i_1, i_2, \dots, i_k)$$

G_μ multiplicitását, tehát azon gyermekek számát, akik rendelkeznek a G_μ -höz tartozó CA-kkal, de más CA-val nem,

$$O_T(G_\mu) = O_T(i_1, i_2, \dots, i_k)$$

pedig azokét, akik rendelkeznek a G_μ -höz tartozó CA-kkal és esetleg továbbiakkal is. Nyilvánvaló a következő összefüggés:

$$O_T(G_\mu) = \sum_{v: G_v \supseteq G_\mu} O(G_v).$$

Az adatok számítógépes kezelését illetően ki kellett elégíteni a következő két követelményt: i) az adatok a memória ne túl nagy részét foglalják el; ii) tetszőleges $1 \leq \mu \leq 2^n - 1$ mellett az $O(G_\mu)$, $O_T(G_\mu)$ számok gyorsan kikereshetőek legyenek. Ezeket a számokat és címüket, illetve az értékükre vonatkozó információt egy egyindexes NS egész tömbben helyeztük el. Ennek feltöltése egy EU változó segítségével történt, amelynek értéke mindig NS első feltöltetlen elemének a sorszáma, az alábbiak szerint:

$$1) \text{ EU} := n + 1;$$

$$2) \text{ } i = 1, 2, \dots, n\text{-re } \text{NS}[i] := \text{EU},$$

$$\text{NS}[\text{EU}] := \text{NS}[\text{EU} + 1] := O(i), \text{ és}$$

$$\text{EU} := \text{EU} + 2 + n - i;$$

3. $G_\mu = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$, $i_1 < i_2 < \dots < i_k$, $k \geq 2$, $O(G_\mu) > 0$, esetén a

$$G_v = \{A_{i_1}\}, \{A_{i_2}\}, \dots, \{A_{i_k}\}, \{A_{i_1}, A_{i_2}\}, \\ \{A_{i_1}, A_{i_3}\}, \dots, \{A_{i_{k-1}}, A_{i_k}\}, \dots, \{A_{j_1}, A_{j_2}, \dots, A_{j_r}\}, \dots, G_\mu$$

részalmazatokra ($r \leq k$, $j_1 < j_2 < \dots < j_r \in \{i_1, i_2, \dots, i_k\}$)

$$S_1 := j_1, \quad S_2 := NS[S_1] + 1 + j_2 - j_1, \dots, S_r := NS[S_{r-1}] + 1 + j_r - j_{r-1},$$

$NS[S_r] = 0$ esetén $NS[S_r] := EU$ és $EU := EU + 2 + n - j_r$, továbbá

$$NS[NS[S_r]] = O_T(G_v) := NS[NS[S_r]] + O(G_\mu),$$

és $G_v = G_\mu$ esetén $NS[NS[S_r] + 1] := O(G_\mu)$.

4.1. TÉTEL. *A fentiek szerint feltöltött NS tömbből tetszőleges*

$$G_\mu = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}, \quad 1 \leq k \leq n, \quad 1 \leq \mu \leq 2^n - 1,$$

esetén az $O_T(G_\mu)$, $O(G_\mu)$ számok legfeljebb $(k+1)$ számú lépésben kikereshetők.

Bizonyítás. $T_1 := NS[i_1] + 1 + i_2 - i_1$, $T_j := NS[T_{j-1}] + 1 + i_{j+1} - i_j$, $j = 2, 3, \dots, k-1$, $T_k := NS[T_{k-1}]$. Ha bármelyik T_j ($1 \leq j \leq k$) értéke 0, akkor $O_T(G_\mu) = O(G_\mu) = 0$, egyébként pedig $O_T(G_\mu) = NS[T_k]$ és $O(G_\mu) = NS[T_k + 1]$. Ezzel a tétellel már adódik.

NS méretét háromféleképpen is csökkentettük:

1. a memóriában az A_i rendellenességeket $O(i)$ csökkenő sorrendjében rendeztük;
2. figyelembe véve az $O_T(G_\mu)$ számok várható nagyságrendjét, NS minden 4 elemébe 5 értéket helyeztünk el;
3. az előfordult sok rendellenességű G_μ -khöz, mivel ilyen kevés volt, manuálisan határoztuk meg az $O_T(G_\mu)$, $O(G_\mu)$ számokat.

Ezen csökkentések után az előfordult 921-féle G_μ -höz egy 9268 elemű NS volt szükséges.

5. Az MCA-k méretének eloszlása

Tetszőleges G_μ esetén jelölje $|G_\mu|$ G_μ méretét, vagyis a G_μ -höz tartozó CA-k számát. Legyen

$$K = \max_\mu |G_\mu|$$

(esetünkben 7), jelölje $O^{(k)}$ azon gyermekek számát, akik pontosan k számú CA-val rendelkeznek ($k = 0, 1, \dots, K$), és legyen $O = O^{(0)}$, akkor

$$\sum_{k=2}^K O^{(k)} = 2762,$$

$$O^{(1)} = 30850 - 2762 = 28088,$$

$$O = M - 30850 = 115926.$$

Az egyes $O^{(k)}$ -k értéke $k \geq 2$ esetén a következő:

$$O^{(2)} = 2128,$$

$$O^{(3)} = 448,$$

$$O^{(4)} = 122,$$

$$O^{(5)} = 50,$$

$$O^{(6)} = 11,$$

$$O^{(7)} = 3.$$

$k=1, 2, \dots, K$ és $r=1, 2, \dots, k$ mellett a pontosan k számú CA-val rendelkező gyermekek körében előfordult pontosan r elemű CA-kombinációk száma $\binom{k}{r} O^{(k)}$, az összes (M számú) gyermek körében előfordult pontosan r elemű CA-kombinációké pedig

$$(5.1) \quad \sum_{k=r}^K \binom{k}{r} O^{(k)}$$

($r=1, 2, \dots, K$). (5.1) értéke $r=1$ esetén 34 513: a 30 850 CA-s gyermek összesen ennyi CA-val rendelkezett. Ez azt jelenti, hogy amíg egy tetszőleges gyermek átlagosan

$$34\,513/M = 0,0291,$$

addig egy CA-s gyermek átlagosan

$$34\,513/30\,850 = 1,119$$

CA-val rendelkezett. (5.1) értéke $r=2, 3, \dots, 7$ esetén rendre 4932, 1761, 642, 179, 32 és 3.

$k=0, 1, \dots, K$ mellett az elméletileg lehetséges különböző k elemű CA-kombinációk száma $\binom{n}{k}$. Jelölje $N_T^{(k)}$ az előfordultak számát, $N^{(k)}$ pedig ezek közül azokét, amelyek a pontosan k számú CA-val rendelkező gyermekek körében fordultak elő. Nyilvánvaló módon

$$N^{(k)} \leq N_T^{(k)} \leq \binom{n}{k}.$$

Mivel a CA-k között vannak egymást definíció szerint kizáróak, azért

$$N_T^{(k)} < \binom{n}{k}.$$

$k=0, 1, \dots, K$ esetén $N^{(k)}$ és $N_T^{(k)}$ értéke a következő:

$$N^{(0)} = N_T^{(0)} = 1,$$

$$N^{(1)} = N_T^{(1)} = n = 40,$$

$$N^{(2)} = 358, \quad N_T^{(2)} = 562, \quad \left[\binom{n}{2} = 780, \right]$$

$$N^{(3)} = 340, \quad N_T^{(3)} = 1141, \quad \left[\binom{n}{3} = 9880, \right]$$

$$N^{(4)} = 119, \quad N_T^{(4)} = 612,$$

$$N^{(5)} = 50, \quad N_T^{(5)} = 178,$$

$$N^{(6)} = 11, \quad N_T^{(6)} = 32,$$

$$N^{(7)} = N_T^{(7)} = 3.$$

Legyen

$$N = \sum_{k=2}^K N^{(k)}$$

az előfordult különböző MCA-k száma, akkor $N=881$, és

$$\sum_{k=0}^K N^{(k)} = N^{(0)} + N^{(1)} + N = 1 + 921 = 922.$$

Ezen kívül

$$\sum_{k=0}^K N_T^{(k)} = 2569.$$

6. A függetlenségi koncepció

Ugyanaz a CA különböző esetekben különböző kóreredetű lehet (a kóreredet általában nincs bejelentve, ezért az MCA-k statisztikai vizsgálatánál ismeretlen volt). Általánosan elfogadott, hogy a CA-s gyermekek jelentős részénél a CA(k) multifaktoriális kóreredetű(ek, lásd [8]). Az MCA-k egy kézenfekvő és sokáig elfogadott magyarázatát a CA-k ún. „függetlenségi koncepciója” nyújtja. Ennek lényege, hogy a multifaktoriális kóreredet szoros kapcsolatban van a CA-k véletlen előfordulásával. Egy multifaktoriális kóreredetű CA független a többi CA-tól, az ilyen CA-t tartalmazó MCA-k tehát véletlen kombinációk. A többi MCA (amelyeket ezért szisztematikus MCA-knak is nevezhetünk) monofaktoriális kóreredetű. A két típus azonban — a kóreredet bejelentetlensége miatt — vizsgálatunkban megkülönböztethetetlen.

A függetlenségi koncepció nem teszi fel a CA-k bármilyen kóreredet melletti függetlenségét, tehát azt, hogy minden MCA véletlen kombináció. Ez utóbbi hipotézist egy viszonylag kis, 232 MCA-s gyermeket tartalmazó vizsgálati anyag alapján [6] elutasította, mivel két vagy több CA ugyanannál a gyermeknél gyakrabban fordult elő, mint az függetlenség esetén várható lett volna. Lényegesnek látszott azonban

ennek az elutasításnak a jogosságát a fentebb ismertetett hétéves, országos anyagon ellenőrizni. Ezen túlmenően fontos jellemezni az MCA-k és a véletlen kombinációk gyakorisága eltéréseinek mértékét, és ennek fényében dönteni a függetlenségi koncepció elfogadásáról.

Tetszőleges

$$G_\mu = \{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}, \quad 1 \leq k \leq n, \quad 1 \leq \mu \leq 2^n - 1,$$

esetén jelölje $E(i_1, i_2, \dots, i_k)$ az olyan gyermekek várható számát, akik rendelkeznek a G_μ -höz tartozó CA-kkal, de más CA-val nem, $P_T(i_1, i_2, \dots, i_k)$ annak a valószínűségét, hogy egy gyermek rendelkezik a G_μ -höz tartozó CA-kkal és esetleg továbbiakkal is, $E_T(i_1, i_2, \dots, i_k)$ pedig az ilyen gyermekek várható számát. Jelölje $P^{(k)}$ annak a valószínűségét, hogy egy gyermek pontosan k számú CA-val rendelkezik, $E^{(k)}$ az ilyen gyermekek várható számát ($k=0, 1, \dots, n$), továbbá legyen $P=P^{(0)}$ és $E=E^{(0)}$. $k=1, 2, \dots, n$, valamint $j=1, 2, \dots, k$ és $1 \leq i_1 < i_2 < \dots < i_j \leq n$ esetén jelölje $P^{(k)}(i_1, i_2, \dots, i_j)$ annak a valószínűségét, hogy egy gyermek pontosan k számú CA-val rendelkezik, köztük A_{i_1} -gyel, A_{i_2} -vel, \dots , A_{i_j} -vel, $O^{(k)}(i_1, i_2, \dots, i_j)$ pedig az ilyen gyermekek számát. Vezessük még be a következő jelöléseket:

$$q_i = \frac{P_T(i)}{1 - P_T(i)}, \quad i = 1, 2, \dots, n,$$

$$S_0 = 1, \quad S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \prod_{j=1}^k q_{i_j}, \quad k = 1, 2, \dots, n,$$

$$\bar{S}_k(i) = \sum_{\substack{i \leq i_1 < i_2 < \dots < i_k \leq n \\ i_1, i_2, \dots, i_k \neq i}} \prod_{j=1}^k q_{i_j}, \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, n-1.$$

Tegyük most fel a CA-k bármilyen kóreredit melletti függetlenségét. Ezen feltétel mellett tetszőleges $k=1, 2, \dots, n$ és $1 \leq i_1 < i_2 < \dots < i_k \leq n$ esetén

$$P_T(i_1, i_2, \dots, i_k) = \prod_{j=1}^k P_T(i_j).$$

$P^{(k)}$, $P^{(k)}(i)$, S_k és $\bar{S}_k(i)$ definíciójából egyszerű számolással következik, hogy

$$P^{(k)} = PS_k, \quad k = 0, 1, \dots, n,$$

$$P^{(k)}(i) = Pq_i \bar{S}_{k-1}(i), \quad i = 1, 2, \dots, n, \quad k = 2, 3, \dots, n,$$

$$\bar{S}_j(i) = S_j - q_i \bar{S}_{j-1}(i), \quad i = 1, 2, \dots, n, \quad j = 2, 3, \dots, n-1,$$

és

$$\bar{S}_1(i) = S_1 - q_i,$$

tehát az $O^{(k)}/M$ és $O^{(k)}(i)/M$ relatív gyakoriságoknak megfelelő $P^{(k)}$ és $P^{(k)}(i)$ (elméleti) valószínűségek előállításához elég az S_k -kat kiszámítanunk, amelyek — a q_i -ken keresztül — csak a $P_T(i)$ valószínűségektől függenek.

6.1. TÉTEL. $k=1, 2, \dots, n$ esetén

$$S_k = \frac{1}{k} \sum_{j=1}^k (-1)^{j-1} T_j S_{k-j},$$

ahol

$$T_j = \sum_{i=1}^n q_i^j, \quad j = 1, 2, \dots, n.$$

Bizonyítás. $S_1 = T_1 = T_1 S_0$.

Legyen

$$U_m^{(j)} = \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} \sum_{v=1}^m \left(q_{i_v}^{j-1} \prod_{\mu=1}^m q_{i_\mu} \right), \quad m, j = 1, 2, \dots, n,$$

akkor

$$U_m^{(1)} = m S_m, \quad m = 1, 2, \dots, n,$$

$$U_1^{(j)} = T_j, \quad j = 1, 2, \dots, n,$$

továbbá

$$U_m^{(j)} + U_{m-1}^{(j+1)} = T_j S_{m-1}, \quad m = 2, 3, \dots, n, \quad j = 1, 2, \dots, n-1.$$

Innen

$$\begin{aligned} k S_k &= U_k^{(1)} = T_1 S_{k-1} - U_{k-1}^{(2)} = \dots = \sum_{j=1}^{k-1} (-1)^{j-1} T_j S_{k-j} + (-1)^{k-1} U_1^{(k)} = \\ &= \sum_{j=1}^{k-1} (-1)^{j-1} T_j S_{k-j} + (-1)^{k-1} T_k = \sum_{j=1}^{k-1} (-1)^{j-1} T_j S_{k-j} + (-1)^{k-1} T_k S_0 = \\ &= \sum_{j=1}^k (-1)^{j-1} T_j S_{k-j}, \quad k = 2, 3, \dots, n. \end{aligned}$$

Ezzel a tétel már adódik.

Mint az előző fejezetben láttuk, az MCA-k mérete tapasztalati eloszlásának jellemzésében az $N^{(k)}$ és $N_T^{(k)}$ mennyiségek is fontos szerepet játszanak. Ezek elméleti megfelelője az $\{O(i_1, i_2, \dots, i_k) > 0\}$, ill. az $\{O_T(i_1, i_2, \dots, i_k) > 0\}$ események valószínűségeinek $1 \leq i_1 < i_2 < \dots < i_k \leq n$ mellett vett összege. Az $O(i_1, i_2, \dots, i_k)$, ill. $O_T(i_1, i_2, \dots, i_k)$ változók binomiális eloszlását $E(i_1, i_2, \dots, i_k)$, ill. $E_T(i_1, i_2, \dots, i_k)$ paraméterű Poisson-eloszlással közelítve a két összeg közelítőleg

$$\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \{1 - \exp[-E(i_1, i_2, \dots, i_k)]\},$$

ill.

$$\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \{1 - \exp[-E_T(i_1, i_2, \dots, i_k)]\}$$

lesz. [Ez abból következik, hogy tetszőleges, λ paraméterű Poisson-eloszlást követő ξ valószínűségi változó esetén

$$P(\xi > 0) = 1 - \exp(-\lambda).]$$

Mivel tetszőleges $1 \leq i_1 < i_2 < \dots < i_k \leq n$, $1 \leq k \leq n$ esetén egyrészt

$$E(i_1, i_2, \dots, i_k) = MP \prod_{j=1}^k q_{i_j},$$

másrészt

$$E_T(i_1, i_2, \dots, i_k) = M \prod_{j=1}^k P_T(i_j),$$

tehát az $N^{(k)}$ és $N_T^{(k)}$ mennyiségek elméleti megfelelőjének közelítése is csak a $P_T(i)$ valószínűségektől függ.

Becsüljük most a $P_T(i)$ valószínűségeket ($i=1, 2, \dots, n$) úgy, hogy $O=E$ és

$$O(i) = E(i), \quad i = 1, 2, \dots, n,$$

teljesüljön. Ekkor

$$\frac{O(i)}{O} = \frac{E(i)}{E}, \quad i = 1, 2, \dots, n.$$

Mivel $E=MP$ és

$$E(i) = MPq_i, \quad i = 1, 2, \dots, n,$$

azért

$$\frac{E(i)}{E} = q_i = \frac{P_T(i)}{1 - P_T(i)}, \quad i = 1, 2, \dots, n,$$

tehát

$$\frac{O(i)}{O} = \frac{P_T(i)}{1 - P_T(i)}, \quad i = 1, 2, \dots, n,$$

ahonnan

$$P_T(i) = \frac{O(i)}{O + O(i)}, \quad i = 1, 2, \dots, n.$$

Mivel

$$E = MP = M \prod_{i=1}^n [1 - P_T(i)]$$

általában nem egyenlő O -val, válasszuk M^* -ot úgy, hogy $E = M^*P$ mellett már teljesüljön $O=E$. Legyen

$$E^{(k)} = M^* P^{(k)}, \quad k = 1, 2, \dots, n,$$

továbbá tetszőleges $1 \leq i_1 < i_2 < \dots < i_k \leq n$, $1 \leq k \leq n$ esetén

$$E(i_1, i_2, \dots, i_k) = M^* P \prod_{j=1}^k q_{i_j}$$

és

$$E_T(i_1, i_2, \dots, i_k) = M^* \prod_{j=1}^k P_T(i_j).$$

Így $k > 1$ -re az $E^{(k)}$ -k következő értékeit kapjuk:

$$E^{(2)} = 303,78,$$

$$E^{(3)} = 1,97,$$

$$E^{(4)} = 8,7 \cdot 10^{-3},$$

$$E^{(5)} = 2,8 \cdot 10^{-5},$$

$$E^{(6)} = 6,8 \cdot 10^{-8},$$

$$E^{(7)} = 1,3 \cdot 10^{-10}.$$

Ezek az $E^{(k)}$ -k igen nagy mértékben különböznek a megfelelő $O^{(k)}$ gyakoriságoktól, ami azt mutatja, hogy a CA-k bármilyen kóreredit melletti függetlensége igen nagy mértékben nem teljesül. (Ez a mérték azonban erősen függ az MCA-mérettől: $E^{(k)}$ és $[O^{(k)} - E^{(k)}]$ aránya $k=2, 3, 4$ mellett rendre 1:6, 1:226 és 1:1,4 · 10⁴, továbbá

$$\sum_{k=5}^7 E^{(k)} : \sum_{k=5}^7 [O^{(k)} - E^{(k)}] = 1:2,3 \cdot 10^6.)$$

$k=2, 3$ mellett $N^{(k)}$ és $N_T^{(k)}$ becslése a következő:

$$\hat{N}^{(2)} = 130,80, \quad \hat{N}_T^{(2)} = 132,57,$$

$$\hat{N}^{(3)} = 1,96, \quad \hat{N}_T^{(3)} = 1,99.$$

Mindenképpen indokolt tehát a CA-k bármilyen kóreredit melletti függetlenségének elutasítása. Ugyanez azonban a függetlenségi koncepcióra is érvényes: ennek fennállása esetén ugyanis az MCA-k 89 százaléka szisztematikus MCA lenne, ami irreális.

Felmerülhet az a gondolat, hogy a CA-k bármilyen kóreredit melletti függetlenségének, ill. a függetlenségi koncepciónak az elfogadhatatlanságát csak néhány CA okozza. Ezért megvizsgáltuk, hogyan változik az egyes CA-k adott méretű MCA-k közti gyakorisága a méret változásával. A 4. táblázat i -edik sora tartalmazza az $O^{(1)}(i), O^{(2)}(i), \dots, O^{(K)}(i)$ és $O_T(i)$ értékeket ($i=1, 2, \dots, n$). $i=1, 2, \dots, n$ mellett az 5. táblázat i -edik sorának k -adik eleme a megfelelő

$$p_i(k) = 100 \frac{O^{(k)}(i)}{\sum_{i=1}^n O^{(k)}(i)}$$

százalékos relatív gyakoriság (annak relatív gyakorisága, hogy egy k számú CA-val rendelkező gyermek rendelkezik A_i -vel, éppen

$$kp_i(k)/100,$$

$k=1, 2, \dots, K$). Az i -edik sor következő, utolsó három eleme közül az első a $p_i(k)$ — k

$$m_i = \frac{\sum_{k=1}^K O^{(k)}(i) p_i(k)}{\sum_{k=1}^K O^{(k)}(i)}$$

súlyozott átlaga, a második az $O^{(k)}(i)$ multiplicitású $p_i(k)$ pontokhoz ($k=1, 2, \dots, K$) a legkisebb négyzetek módszerével illesztett

$$r_i(k) = a_i k + b_i$$

egyenes a_i meredeksége, a harmadik pedig az aszimptotikusan 1 szabadságfokú χ^2 -eloszlást követő

$$(6.1) \quad \sum_{k=1}^K O^{(k)}(i) \left\{ \frac{[p_i(k) - m_i]^2}{m_i} - \frac{[p_i(k) - r_i(k)]^2}{r_i(k)} \right\}$$

4. TÁBLÁZAT

A_i	$O^{(1)}(i)$	$O^{(2)}(i)$	$O^{(3)}(i)$	$O^{(4)}(i)$	$O^{(5)}(i)$	$O^{(6)}(i)$	$O^{(7)}(i)$	$O_T(i)$
AN	984	107	28	2	2	1	0	1124
EN	266	46	21	1	3	1	0	338
SB	1193	124	35	12	2	0	0	1366
CL	1091	172	86	20	13	3	1	1386
CP	362	66	28	19	13	3	1	492
LR	388	63	38	8	9	2	0	508
PY	680	94	52	25	9	4	1	865
SY	291	65	30	6	5	0	1	398
EX	222	75	35	10	5	1	0	348
OA	199	50	21	9	7	3	0	289
AA	210	72	40	16	13	7	3	361
AM	33	25	14	9	4	0	0	85
CT	47	19	4	3	2	0	0	75
HS	1468	163	39	9	4	1	1	1685
EG	43	23	16	10	7	3	2	104
HD	4255	582	183	67	26	5	3	5121
DI	229	67	28	11	5	1	1	342
RA	59	38	22	8	2	2	0	131
CK	97	44	30	9	7	1	1	189
AI	193	28	4	2	1	1	0	229
MC	146	62	30	15	7	0	0	260
HY	750	131	50	27	11	4	1	974
ON	85	1	3	2	0	0	0	91
FN	67	34	30	16	11	3	1	162
EY	146	23	7	8	3	0	0	187
EA	225	45	36	27	17	1	0	351
TC	146	211	35	5	2	0	0	399
BR	62	1	1	1	0	0	0	65
CD	7035	638	68	12	6	0	1	7760
CF	1955	435	101	28	16	4	3	2542
OL	91	27	8	3	1	1	0	131
RS	181	49	19	11	0	2	0	262
PS	395	22	4	0	0	1	0	422
OD	363	55	23	9	2	1	0	453
UT	689	146	42	19	12	3	0	911
OU	390	90	47	17	6	5	0	555
IH	2573	247	37	9	3	0	0	2869
MS	133	66	33	12	11	1	0	256
EO	142	31	10	6	2	0	0	191
ST	204	19	6	5	1	1	0	236

5. TÁBLÁZAT

A_i	$p_i(1)$	$p_i(2)$	$p_i(3)$	$p_i(4)$	$p_i(5)$	$p_i(6)$	$p_i(7)$	m_i	a_i	χ^2
AN	3,50	2,51	2,08	0,41	0,80	1,52	0,00	3,36	-0,79	0,57
EN	0,95	1,08	1,56	0,20	1,20	1,52	0,00	1,01	0,16	0,06
SB	4,25	2,91	2,60	2,46	0,80	0,00	0,00	4,07	-0,91	0,68
CL	3,88	4,04	6,40	4,10	5,20	4,55	4,76	4,07	0,55	0,76
CP	1,29	1,55	2,08	3,89	5,20	4,55	4,76	1,60	0,75	2,00
LR	1,38	1,48	2,83	1,64	3,60	3,03	0,00	1,55	0,46	0,62
PY	2,42	2,21	3,87	5,12	3,60	6,06	4,76	2,59	0,60	1,02
SY	1,04	1,53	2,23	1,23	2,00	0,00	4,76	1,23	0,41	0,44
EX	0,79	1,76	2,60	2,05	2,00	1,52	0,00	1,24	0,60	0,96
OA	0,71	1,17	1,56	1,84	2,80	4,55	0,00	0,98	0,52	0,91
AA	0,75	1,69	2,98	3,28	5,20	10,61	14,29	1,76	1,44	7,95
AM	0,12	0,59	1,04	1,84	1,60	0,00	0,00	0,66	0,47	0,42
CT	0,17	0,45	0,30	0,61	0,80	0,00	0,00	0,28	0,15	0,06
HS	5,23	3,83	2,90	1,84	1,60	1,52	4,76	5,01	-1,09	1,58
EG	0,15	0,54	1,19	2,05	2,80	4,55	9,52	1,07	0,91	2,48
HD	15,15	13,67	13,62	13,73	10,40	7,58	14,29	14,88	-0,92	1,20
DI	0,82	1,57	2,08	2,25	2,00	1,52	4,76	1,15	0,50	0,74
RA	0,21	0,89	1,64	1,64	0,80	3,03	0,00	0,79	0,52	0,60
CK	0,35	1,03	2,23	1,84	2,80	1,52	4,76	0,99	0,66	1,23
AI	0,69	0,66	0,30	0,41	0,40	1,52	0,00	0,68	-0,02	0,00
MC	0,52	1,46	2,23	3,07	2,80	0,00	0,00	1,15	0,77	1,47
HY	2,67	3,08	3,72	5,53	4,40	6,06	4,76	2,89	0,63	1,05
ON	0,30	0,02	0,22	0,41	0,00	0,00	0,00	0,30	0,00	0,00
FN	0,24	0,80	2,23	3,28	4,40	4,55	4,76	1,41	0,98	2,19
EY	0,52	0,54	0,52	1,64	1,20	0,00	0,00	0,58	0,22	0,14
EA	0,80	1,06	2,68	5,53	6,80	1,52	0,00	1,68	1,35	5,94
TC	0,52	4,96	2,60	1,02	0,80	0,00	0,00	3,06	1,64	2,15
BR	0,22	0,02	0,07	0,20	0,00	0,00	0,00	0,21	-0,04	0,00
CD	25,05	14,99	5,06	2,46	2,40	0,00	4,76	23,99	-9,18	40,71
CF	6,96	10,22	7,51	5,74	6,40	6,06	14,29	7,53	0,80	1,60
OL	0,32	0,63	0,60	0,61	0,40	1,52	0,00	0,42	0,16	0,07
RS	0,64	1,15	1,41	2,25	0,00	3,03	0,00	0,88	0,48	0,55
PS	1,41	0,52	0,30	0,00	0,00	1,52	0,00	1,35	-0,43	0,18
OD	1,29	1,29	1,71	1,84	0,80	1,52	0,00	1,32	0,11	0,03
UT	2,45	3,43	3,13	3,89	4,80	4,55	0,00	2,71	0,55	0,70
OU	1,39	2,11	3,50	3,48	2,40	7,58	0,00	1,82	0,84	2,21
IH	9,16	5,80	2,75	1,84	1,20	0,00	0,00	8,76	-3,03	0,00
MS	0,47	1,55	2,46	2,46	4,40	1,52	0,00	1,27	0,86	1,98
EO	0,51	0,73	0,74	1,23	0,80	0,00	0,00	0,58	0,17	0,07
ST	0,73	0,45	0,45	1,02	0,40	1,52	0,00	0,71	0,00	0,00

kifejezés értéke ($i=1, 2, \dots, n$). a_i azt mutatja, hogyan változik a $p_i(k)$ százalékos relatív gyakoriság az A_i -t tartalmazó CA-kombináció méretének növekedésével, (6.1) pedig azt, hogy mennyivel becsüli $r_i(k)$ jobban $p_i(k)$ -t, mint m_i ($i=1, 2, \dots, n$). (6.1) csupán három CA mellett szignifikáns: AA és EA esetén, amikor az a_i meredekség pozitív, valamint CD esetén, amikor a_i negatív. Tehát az adott méretű MCA-k közti egyes CA-gyakoriságok a méret változásával lényegében egymáshoz hasonlóan változnak. Ez azt mutatja, hogy a CA-k bármilyen kóreredet melletti függetlenségének, ill. a függetlenségi koncepciónak az elfogadhatatlanságát nem csupán néhány CA okozza.

7. A függetlenségi koncepció módosítása

A függetlenségi koncepció „duálisa” az a — lényegében [5]-ből származó — feltételezés, amely szerint a szisztematikus MCA-kban a CA-k függetlenek, de a multifaktoriális kóreredetűek nem feltétlen azok. $k=1, 2, \dots, n$, valamint $j=1, 2, \dots, k$ és $1 \leq i_1 < i_2 < \dots < i_j \leq n$ esetén jelölje $E^{(k)}(i_1, i_2, \dots, i_j)$ az olyan gyermekek várható számát, akik pontosan k számú CA-val rendelkeznek, köztük A_{i_1} -gyel, A_{i_2} -vel, ..., A_{i_j} -vel. Szem előtt tartva, hogy egyrészt a nagyméretű MCA-k még a multifaktoriális kóreredetű CA-k függősége esetén is — orvosi megfontolások alapján — zömmel szisztematikusak, másrészt a legalább 5 CA-val rendelkező gyermekek

$$\sum_{k \geq 5} O^{(k)} = O^{(5)} + O^{(6)} + O^{(7)}$$

száma csupán 64, a duális koncepció ellenőrzéséhez a $P_T(i)$ valószínűségeket ($i=1, 2, \dots, n$) úgy próbáljuk becsülni, hogy $O^{(3)} = E^{(3)}$ és

$$(7.1) \quad O^{(4)}(i) = E^{(4)}(i), \quad i = 1, 2, \dots, n,$$

teljesüljön. Mivel

$$E^{(3)} = MP^{(3)} = MPS_3,$$

és

$$E^{(4)}(i) = MP^{(4)}(i) = MPq_i \bar{S}_3(i), \quad i = 1, 2, \dots, n,$$

azért

$$\frac{E^{(4)}(i)}{E^{(3)}} = q_i \frac{\bar{S}_3(i)}{S_3} = \frac{P_T(i)}{1 - P_T(i)} \cdot \frac{\bar{S}_3(i)}{S_3}, \quad i = 1, 2, \dots, n,$$

tehát

$$\frac{O^{(4)}(i)}{O^{(3)}} = \frac{P_T(i)}{1 - P_T(i)} \cdot \frac{\bar{S}_3(i)}{S_3}, \quad i = 1, 2, \dots, n,$$

ahonnan egyszerű számolással adódik, hogy

$$P_T(i) = \frac{O^{(4)}(i)}{O^{(4)}(i) + MPS_3(i)}, \quad i = 1, 2, \dots, n.$$

Mivel

$$E^{(3)} = MPS_3 = MS_3 \prod_{i=1}^n [1 - P_T(i)]$$

általában nem egyenlő $O^{(3)}$ -mal, válasszuk \tilde{M} -ot és $P_T(1), P_T(2), \dots, P_T(n)$ becslését úgy, hogy $E^{(3)} = \tilde{M}P^{(3)}$ és

$$E^{(4)}(i) = \tilde{M}P^{(4)}(i), \quad i = 1, 2, \dots, n,$$

mellett teljesüljön $O^{(3)} = E^{(3)}$ és (7.1). Ezt a következő iteráció segítségével érjük el:

$$\begin{aligned} \tilde{M}_0 &= M - M^*, \quad P_{T,0}(i) = \frac{O_T(i) - O(i)}{\tilde{M}_0}, \quad i = 1, 2, \dots, n, \\ \tilde{M}_{m+1} &= \frac{O^{(3)}}{P_m S_{3,m}}, \quad P_{T,m+1}(i) = \frac{O^{(4)}(i)}{O^{(4)}(i) + \tilde{M}_{m+1} P_m \bar{S}_{3,m}(i)}, \\ &i = 1, 2, \dots, n; \quad m = 1, 2, \dots \end{aligned}$$

Legyen

$$E^{(k)} = \tilde{M}P^{(k)}, \quad k = 1, 2, \dots, n,$$

továbbá tetszőleges $1 \leq i_1 < i_2 < \dots < i_k \leq n$, $1 \leq k \leq n$ esetén

$$E(i_1, i_2, \dots, i_k) = \tilde{M}P \prod_{j=1}^k q_{i_j}$$

és

$$E_T(i_1, i_2, \dots, i_k) = \tilde{M} \prod_{j=1}^k P_T(i_j).$$

Így $k \neq 3, 4$ mellett az $E^{(k)}$ -k következő értékeit kapjuk:

$$E^{(0)} = 1426,$$

$$E^{(1)} = 1889,$$

$$E^{(2)} = 1163,$$

$$E^{(5)} = 25,$$

$$E^{(6)} = 4,$$

$$E^{(7)} = 1.$$

Mivel az $E^{(0)}$, $E^{(1)}$, $E^{(2)}$ értékek itt nem különösebben érdekesek — hiszen a szisztematikus MCA-k általában nagyobb méretűek —, a duális koncepció ránézésre nem elfogadhatatlan. Ahhoz azonban, hogy használható is legyen, feltétlenül ki kell egészíteni a multifaktoriális kóreredetű CA-kra vonatkozó valamilyen feltevessel. Ezt teszi a feltételes függetlenségi koncepció, amely a függetlenségi koncepciónak és duálisának mintegy „metszete”. Eszerint tetszőleges MCA p_1 , ill. p_2 valószínűséggel multifaktoriális, ill. monofaktoriális kóreredetű, és a CA-k a kóreredet mint feltétel mellett függetlenek. Ez azt jelenti, hogy az MCA-k valószínűségeloszlása $m=2$ számú olyan eloszlás keveréke, amelyek mindegyikében a CA-k függetlenek. Mivel a függetlenségi koncepció elfogadhatatlan volt, nyilván a feltételes függetlenségi koncepció is az. Viszont természetes módon általánosítható: nem tesszük fel, hogy akár a multifaktoriális, akár a monofaktoriális kóreredetű MCA-khoz a keverék egyetlen eloszlása tartozik, tehát azt, hogy $m=2$. Az így konstruált keverékeloszlásos modell [7]-ben és [4]-ben kerül kifejtésre.

Köszönetet mondok CZEIZEL ENDRÉNEK és TUSNÁDY GÁBORNAK értékes tanácsaikért és a dolgozat megírásában nyújtott segítségükért.

IRODALOM

- [1] COHEN, M. M., JR., "On the nature of syndrome delineation", *Acta Genet. Med. Gemellol.* **103** (1977) 103—119.
- [2] CZEIZEL, A., "Definition of multiple congenital abnormalities", *Acta Morph. Acad. Sci. Hung.* **29** (1981) 251—258.
- [3] CZEIZEL, A., PÁZSY, A., TELEGDİ, L. and TUSNÁDY, G., "Classification and registration of multiple congenital abnormalities", *Acta Morph. Acad. Sci. Hung.* **29** (1981) 377—390.
- [4] CZEIZEL, A., TELEGDİ, L. and TUSNÁDY, G., *Multiple Congenital Abnormalities* (Akadémiai Kiadó, Budapest, sajtó alatt).

- [5] KÄLLEN, B. and WINBERG, J., "Multiple malformations studied with a national register of malformations", *Pediatrics* **44** (1969) 410—417.
- [6] ROBERTS, C. J. and POWELL, R. G., "Interrelation of the common congenital malformations. Some aetiological implications", *Lancet* **ii** (1975) 848—854.
- [7] TELEGDI, L., CZEIZEL, A., TUSNÁDY, G., BOLLA, M. and PÁZSY, A., "Statistical study on the multiple congenital abnormalities in Hungary, 1970—1976", Working Paper, MS/6. MTA SZTAKI, Budapest, 1981.
- [8] TUSNÁDY, G., TELEGDI, L. és CZEIZEL, E., „Gyakori veleszületett rendellenességek öröklődésmentének vizsgálata”, *Alk. Mat. Lapok* **4** (1978) 1—25.

(Beérkezett: 1983. május 23.)

TELEGDI LÁSZLÓ

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE
1111 BUDAPEST XI., KENDE U. 13—17.

INVESTIGATION OF THE INDEPENDENCE OF CONGENITAL ABNORMALITIES

L. TELEGDI

The paper deals with the statistical analysis of multiple congenital abnormalities on the basis of the material of the *Hungarian Congenital Malformation Register*, 1970—1976. The author investigates into the independence of the congenital abnormalities by the estimation of various terms characterizing the distribution of size of the multiple congenital abnormalities.

A kiadásért felelős az Akadémiai Kiadó és Nyomda főigazgatója

Műszaki szerkesztő: Sándor István

A kézirat nyomdába érkezett: 1983. július 18. — Terjedelem: 18,90 (A/5 ív)

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámmal kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék mindig az utolsó szakasz kell hogy legyen, és azt nem kell sorszámmal ellátni. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos postai címét. A dolgozatban előforduló képleteket szak-

ként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segédtevételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német, francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozat ábráit és az esetleges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától független, folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként feltüntetett, ábraazonosító sorszámmal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve társszerzők esetén az első szerző neve szerinti alfabetikus sorrendben úgy, hogy külön, de folytatólagos sorszámozású listát alkossanak a latin és a cirill betűs nevű szerzők műveire vonatkozó hivatkozások, és mindkét részben a megfelelő alfabetikus sorrend legyen kialakítva. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., »Über die Theorie der einfachen Ungleichungen«, *Journal für die reine und angewandte Mathematik* 124 (1902) 1—27.
- [2] Kéri, G., „DUALSIMP”, rutin a CDC 3300-as gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertetők 2. 1973. május) 19—20.
- [3] Prékopa, A., „Sztohasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U., „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam—London, 1973) 221—228.
- [5] Zoutendijk, G., *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76—78]. A szerzők a dolgozatukról 100 darab különlenyomatot kapnak, ezek költsége — nyomott oldalanként 25 forint — a szerzői díjat terheli.

TARTALOMJEGYZÉK

<i>Deák István, Hoffer János, Mayer János, Németh Ágoston, Potecz Béla, Prékopa András és Strazicky Beáta</i> : Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével	221
<i>Dus Magdolna</i> : Tranziens kétfázisú folyadékáramlás matematikai modellezésének problémái ...	339
<i>Polner Gábor, Dallmann Géza és Orosz László</i> : A lambda fág szabályozásának modellezése	357
<i>Kertész Viktor</i> : Pozitív definit kvadratikus Ljapunov függvények alkalmazása stabilitási vizsgálatokhoz	375
<i>Rapcsák Tamás</i> : Megjegyzés az egyenlőtlenség feltételek melletti minimalizálásról	387
<i>Dormány Mihály</i> : Egy bikritérium feladat megoldásáról	393
<i>Molnár Sándor és Szidarovszky Ferenc</i> : A Kriging módszer alkalmazása nemnegatív súlyokkal és korlátokkal	405
<i>Telegdi László</i> : Velezületett rendellenességek függetlenségének vizsgálata	421

INDEX

<i>Deák, I., Hoffer, J., Mayer, J., Németh, Á., Potecz, B., Prékopa, A. and Strazicky, B.</i> , A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Dus, M.</i> , Mathematical modelling of transient two-phase flows	339
<i>Polner, G., Dallmann, G. and Orosz, L.</i> , Modelling of the bacteriophage lambda regulation	357
<i>Kertész, V.</i> , Application of positive definite Lyapunov functions to stability investigations	375
<i>Rapcsák, T.</i> , Remarks on minimization problems with inequality constraints	387
<i>Dormány, M.</i> , On solving a bicriterion program	393
<i>Molnár, S. and Szidarovszky, F.</i> , On the Kriging method with nonnegative weights and bounds ...	405
<i>Telegdi, L.</i> , Investigation of the independence of congenital abnormalities	421

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI
TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

FŐSZERKESZTŐ
PRÉKOPA ANDRÁS

FŐSZERKESZTŐ-HELYETTES
ARATÓ MÁTYÁS

A SZERKESZTŐ BIZOTTSÁG TAGJAI

BENCZUR ANDRÁS, CSISZÁR IMRE, FARKAS MIKLÓS, GYIRES BÉLA,
HATVANI LÁSZLÓ, HEPPES ALADÁR, KÁTAI IMRE, KIS OTTÓ,
SARKADI KÁROLY, TANDORI KÁROLY, VARGA LÁSZLÓ,
SZÁNTAI TAMÁS (technikai szerkesztő)

MUNKATÁRSAK

BAJCSAY PÁL, BALLA KATALIN, BÉKÉSSY ANDRÁS, CSÁKI PÉTER,
CSIRIK JÁNOS, DEMETROVICS JÁNOS, DÉNES JÓZSEF, DÖMÖLKI BÁLINT,
ELBERT ÁRPÁD, FORGÓ FERENC, GÉCSEG FERENC, GERGELY JÓZSEF,
GESZTELYI ERNŐ, GYÖRFFY LÁSZLÓ, KLAFSZKY EMIL, KÓSA ANDRÁS,
KOVÁCS LÁSZLÓ BÉLA, LÁSZLÓ ZOLTÁN, MIKOLÁS MIKLÓS,
MOGYORÓDI JÓZSEF, NÉMETH GÉZA, NEMETZ TIBOR, RÉVÉSZ PÁL, RÓZSA PÁL,
STAHL JÁNOS, SZÉP JENŐ, TANKÓ JÓZSEF, TOMKÓ JÓZSEF, TŐKE PÁL,
TUSNÁDY GÁBOR, VINCZE ENDRE

IX. KÖTET

AKADÉMIAI KIADÓ, BUDAPEST
1983

TARTALOMJEGYZÉK

<i>Biró Miklós</i> : A legerősebb próbák témakörének matematikai programozási megközelítése	105
<i>Biró Miklós</i> : A bináris hátizsák feladat	113
<i>Dallmann Géza, Polner Gábor és Orosz László</i> : A lambda fág szabályozásának modellezése . . .	357
<i>Deák István, Hoffer János, Mayer János, Németh Ágoston, Potecz Béla, Prékopa András és Strazicky Beáta</i> : Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével	221
<i>Demetrovics János, Füredi Zoltán és Katona Gyula</i> : A függőségek és az individuumok száma közötti kapcsolat összetett adatrendszerek esetén	13
<i>Dormány Mihály</i> : Egy bikritérium feladat megoldásáról	393
<i>Dus Magdolna</i> : Tranziens kétfázisú folyadékáramlás matematikai modellezésének problémái . .	339
<i>Érdi Péter, Tóth János és Török László Turul</i> : A Poisson-eloszlás jelentősége összetett kémiai reakciók sztochasztikus modelljében	175
<i>Farkas Miklós</i> : Matematika és objektív valóság	1
<i>Fülöp János</i> : Eredeti csúcspont keresése és alkalmazása konkáv függvény lineáris feltételek melletti minimalizálásakor	51
<i>Füredi Zoltán, Demetrovics János és Katona Gyula</i> : A függőségek és az individuumok száma közötti kapcsolat összetett adatrendszerek esetén	13
<i>Füredi Zoltán</i> : Minimális relációs adatbázisok	23
<i>Galántai Aurél</i> : Egylépéses módszercsaládok konvergencia- és hibaanalízise	29
<i>Gyires Béla</i> : Egy mátrixegyenlet megoldásáról és az eredmények alkalmazásáról valószínűségi eloszlásfüggvények lineáris kombinációinak elméletében	137
<i>Hoffer János, Deák István, Mayer János, Németh Ágoston, Potecz Béla, Prékopa András és Strazicky Beáta</i> : Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével	221
<i>Huhn Edit</i> : Kálmán-szűrők aszimptotikus viselkedéséről	197
<i>Katona Gyula, Demetrovics János és Füredi Zoltán</i> : A függőségek és az individuumok száma közötti kapcsolat összetett adatrendszerek esetén	13
<i>Kertész Viktor</i> : Pozitív definit kvadratikus Ljapunov-függvények alkalmazása stabilitási vizsgálatokhoz	375
<i>Komáromi Éva</i> : Logaritmikusan konkáv eloszlásfüggvénnyel korlátozott lineáris programozási feladat megoldásának duális megközelítése	85
<i>Mayer János, Deák István, Hoffer János, Németh Ágoston, Potecz Béla, Prékopa András és Strazicky Beáta</i> : Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével	221
<i>Molnár Sándor és Szidarovszky Ferenc</i> : A Kriging-módszer alkalmazása nemnegatív súlyokkal és korlátokkal	405
<i>Németh Ágoston, Deák István, Hoffer János, Mayer János, Potecz Béla, Prékopa András és Strazicky Beáta</i> : Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével	221
<i>Orosz László, Polner Gábor és Dallmann Géza</i> : A lambda fág szabályozásának modellezése . .	357
<i>Pick Róbert</i> : A maximum likelihood becslés aszimptotikus optimalitásáról	203
<i>Pintér János</i> : A Csebisev-egyenlőtlenség egy élesítése Monte Carlo-módszerrel becsült függvényértékek esetén	93

<i>Polner Gábor, Dallmann Géza és Orosz László: A lambda fág szabályozásának modellezése ...</i>	357
<i>Potecz Béla, Deák István, Hoffer János, Mayer János, Németh Ágoston, Prékopa András és Strazicky Beáta: Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével</i>	221
<i>Prékopa András, Deák István, Hoffer János, Mayer János, Németh Ágoston, Potecz Béla és Strazicky Beáta: Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével</i>	221
<i>Rapcsák Tamás: A nemlineáris programozási feladat optimalitási feltételeinek differenciál-geometriai vizsgálata</i>	73
<i>Rapcsák Tamás: Megjegyzés az egyenlőtlenség feltételek melletti minimalizálásról</i>	387
<i>Sparing László: A perturbált van der Pol-differenciálegyenlet számítógépes vizsgálata</i>	43
<i>Strazicky Beáta, Deák István, Hoffer János, Mayer János, Németh Ágoston, Potecz Béla és Prékopa András: Nagyméretű, vegyesváltozós, matematikai modell termikus erőműves villamosenergia-rendszer rövidtávú, optimális menetrendjének meghatározására hálózati feltételek figyelembevételével</i>	221
<i>Szidarovszky Ferenc és Molnár Sándor: A Kriging-módszer alkalmazása nemnegatív súlyokkal és korlátokkal</i>	405
<i>Telegdi László: Veszélytett rendellenességek függetlenségének vizsgálata</i>	421
<i>Tóth János, Érdi Péter és Török László Turul: A Poisson-eloszlás jelentősége összetett kémiai reakciók sztochasztikus modelljében</i>	175
<i>Török László Turul, Érdi Péter és Tóth János: A Poisson-eloszlás jelentősége összetett kémiai reakciók sztochasztikus modelljében</i>	175
<i>Varga Gyula: Párhuzamos algoritmus polinomok másodfokú tényezőkre bontására</i>	145
<i>Vetier András: Szinaj-biliárd potenciáltérben, horocikluskonstrukció</i>	217
<i>Hírek és közlemények</i>	217

INDEX

<i>Biró, M., On a mathematical programming approach to the topic of the most efficient tests ..</i>	105
<i>Biró, M., The binary knapsack problem</i>	113
<i>Dallmann, D., Polner, G. and Orosz, L., Modelling of the bacteriophage lambda regulations ..</i>	357
<i>Deák, I., Hoffer, J., Mayer, J., Németh, Á., Potecz, B., Prékopa, A. and Strazicky, B., A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account</i>	221
<i>Demetrovics, J., Füredi, Z. and Katona, G. O. H., The relation between the number of individuals and independencies in the composed data bases</i>	13
<i>Dormány, M., On solving a bicriterion program</i>	393
<i>Dus, M., Mathematical modelling of transient two-phase flows</i>	339
<i>Érdi, P., Tóth, J. and Török, T. L., Significance of the Poisson distribution in the stochastic model of complex chemical reactions</i>	175
<i>Farkas, M., Mathematics and objective reality</i>	1
<i>Fülöp, J., Finding original vertex and its application minimizing a concave function subject to linear constraints</i>	51
<i>Füredi, Z., Demetrovics, J. and Katona, G. O. H., The relation between the number of individuals and dependencies in the composed data bases</i>	13
<i>Füredi, Z., Minimal relational data bases</i>	23
<i>Galántai, A., Convergence theorems and error analysis for the families of one-step methods ..</i>	29
<i>Gyires, B., On the solution of certain matrix equation and an application of the results in the theory of linear combinations of probability distribution functions</i>	137
<i>Hoffer, J., Deák, I., Mayer, J., Németh, Á., Potecz, B., Prékopa, A., and Strazicky, B., A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account</i>	221
<i>Huhn, E., On the asymptotic behaviour of Kalman filters</i>	197

<i>Katona, G. O. H., Demetrovics, J., and Füredi, Z.,</i> The relation between the number of individuals and dependencies in the composed data bases	13
<i>Kertész, V.,</i> Application of the positive definit Lyapunof functions to stability investigations ..	375
<i>Komáromi, É.,</i> A dual approach to stochastic linear programming problems constrained by logarithmic concave joint probability distribution function	85
<i>Mayer, J., Deák, I., Hoffer, J., Németh, Á., Potecz, B., Prékopa, A. and Strazicky, B.,</i> A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Molnár, S. and Szidarovszky, F.,</i> On the Kriging method with nonnegativ weights and bounds	405
<i>Németh, Á., Deák, I., Hoffer, J., Mayer, J., Potecz, B., Prékopa, A., and Strazicky, B.,</i> A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Orosz, L., Polner, G. and Dallmann, G.,</i> Modelling of the bacteriophage lambda regulation ...	357
<i>Pick, R.,</i> On the asymptotic optimality of the maximum likelihood estimation	203
<i>Pintér, J.,</i> An improved Chebyshev-inequality for function value estimates by Monte Carlo techniques	93
<i>Polner, G., Dallmann, G. and Orosz, L.,</i> Modelling of the bacteriophage lambda regulation ..	357
<i>Potecz, B., Deák, I., Hoffer, J., Mayer, J., Németh, Á., Prékopa, A. and Strazicky, B.,</i> A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Prékopa, A., Deák, I., Hoffer, J., Mayer, J., Németh, Á., Potecz, B. and Strazicky, B.,</i> A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Rapcsák, T.,</i> The optimality conditions of the nonlinear programming problems and differencial geometry	73
<i>Rapcsák, T.,</i> Remarks on minimization problems with inequality constraints	383
<i>Sparing, L.,</i> Numerical investigation of the perturbed van der Pol's equation	43
<i>Strazicky, B., Deák, I., Hoffer, J., Mayer, J., Németh, Á., Potecz, B. and Prékopa, A.,</i> A large scale, mixed-variable mathematical programming model for the short-range optimal schedule of electric power systems with thermal power plants taking network constraints into account	221
<i>Szidarovszky, F., and Molnár, S.,</i> On the Kriging method with nonnegativ weights and bounds	405
<i>Telegdi, L.,</i> Investigation of the independence of congenital abnormalities	421
<i>Tóth, J., Erdi, P. and Török, T. L.,</i> Significance of het Poisson distribution in the stochastic model of complex chemical reactions	175
<i>Török, T. L., Tóth, J. and Erdi, P.,</i> Significance of the Poisson distribution in the stochastic model of complex chemical reactions	175
<i>Varga, Gy.,</i> Parallel algorithm for decomposition of polynomials into quadratic factors	
<i>Vetier, A.,</i> Sinai-billiard in potencial fields	145
<i>Communications</i>	217